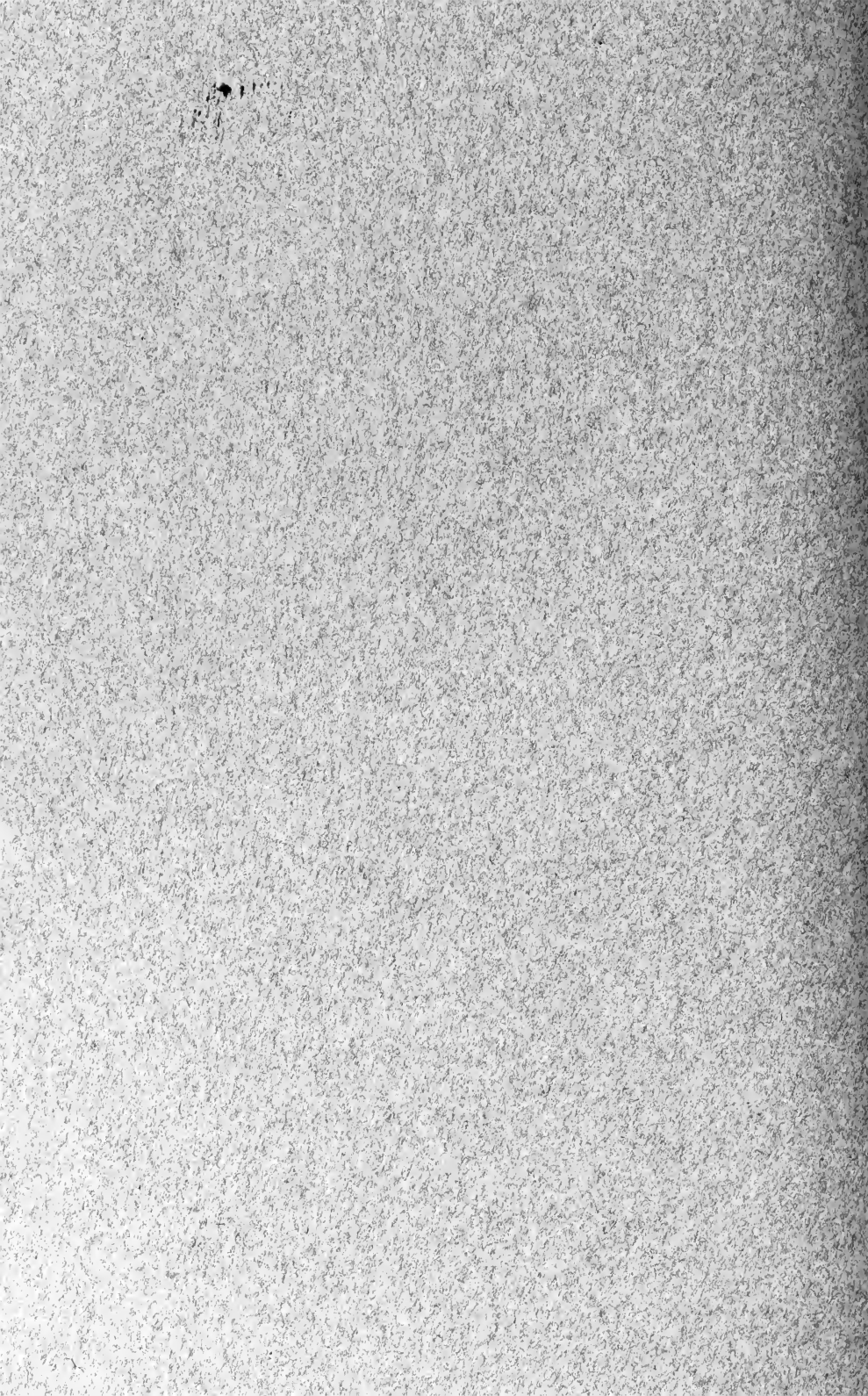


UC-NRLF



B 4 031 751



UNIVERSITY OF ILLINOIS BULLETIN

Vol. XIX

October 31, 1921

No. 9

ISSUED WEEKLY

[Entered as second-class matter December 11, 1912, at the post office at Urbana, Illinois, under the Act of August 24, 1912. Acceptance for mailing at the special rate of postage provided for in section 1103, Act of October 3, 1917, authorized July 31, 1918.]

BUREAU OF EDUCATIONAL RESEARCH—BULLETIN NO. 6

THE ILLINOIS EXAMINATION

BY

Walter S. Monroe

DIRECTOR

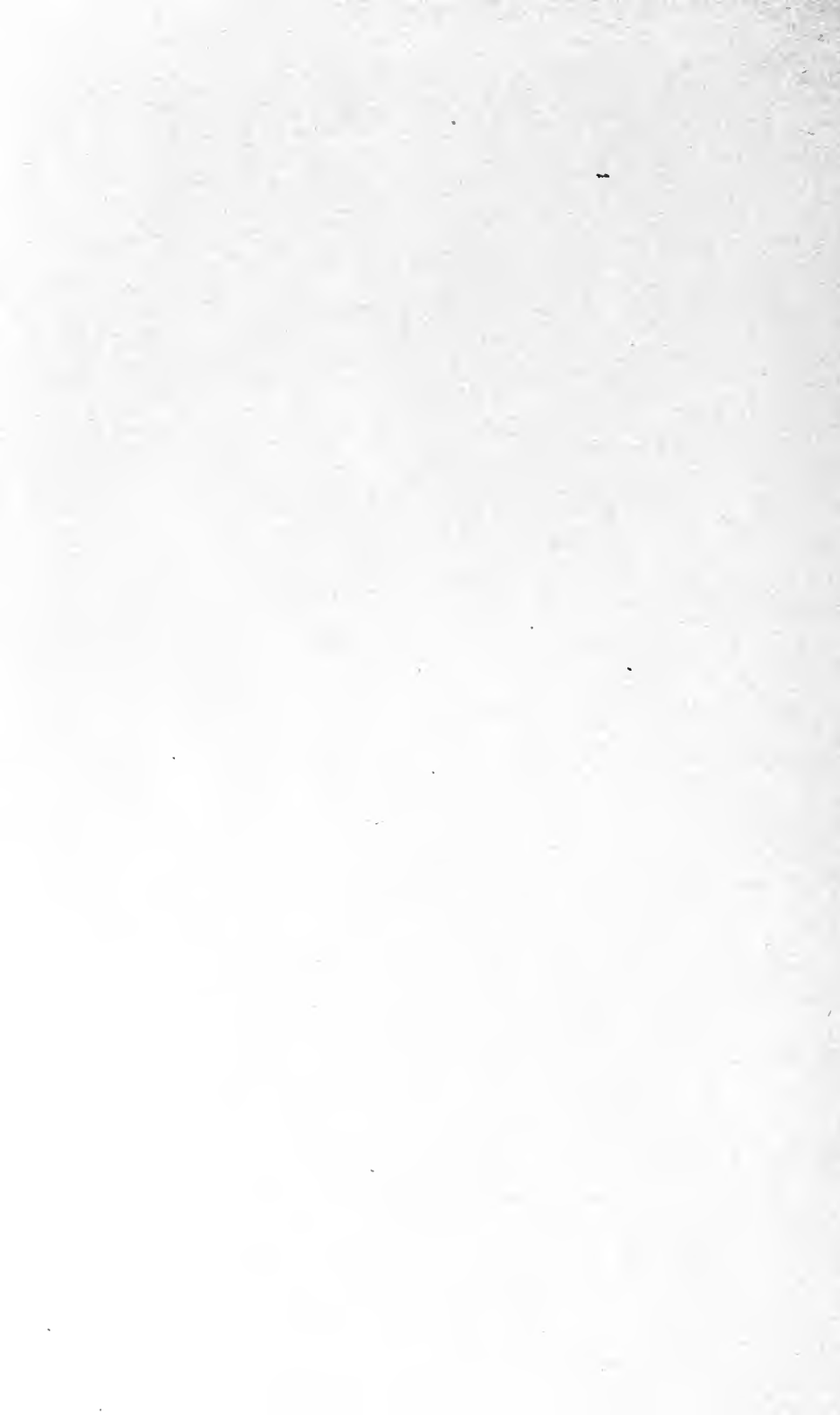
BUREAU OF EDUCATIONAL RESEARCH

COLLEGE OF EDUCATION, UNIVERSITY OF ILLINOIS



PRICE 50 CENTS

PUBLISHED BY THE UNIVERSITY OF ILLINOIS
URBANA



THE ILLINOIS EXAMINATION

BY

Walter S. Monroe

DIRECTOR

Bureau of Educational Research
College of Education, University of Illinois



PUBLISHED BY THE UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS

Digitized by the Internet Archive
in 2007 with funding from
Microsoft Corporation

TABLE OF CONTENTS

	PAGE
PREFACE	4
INTRODUCTION	5
I. FACTS OF TITLE	6
II. NATURE OF PUPIL'S PERFORMANCE	6
1. Illinois General Intelligence Scale	6
2. Monroe's Standardized Reading Tests, Revised.....	10
3. Monroe's General Survey Scale in Arithmetic.....	12
III. DESCRIPTION OF PUPIL'S PERFORMANCE.....	17
1. Point Scores	17
2. Derived Scores	19
IV. FUNCTION	43
V. VALIDITY	43
1. Objectivity	43
2. Reliability	43
3. Discrimination	50
4. Comparison with criterion measures	57
5. Inferences concerning validity based upon the struc- ture of the test and its administration.....	59
VI. VALIDITY OF SIGNIFICANCE	61
VII. NORMS	65

PREFACE.

In this bulletin we present an account of the derivation of the battery of educational tests known as the Illinois Examination. In addition, we include data relative to their validity, reliability, practice effect, norms, and significance. Doubtless some readers will not find all of the information relative to these tests which they desire. Although an effort has been made to have the account complete the author himself is conscious of a number of limitations of this report. It is, however, presented in hopes that it will be of service to those who are interested in the Illinois Examination.

On April 2, 1920, a "Committee on Standard Tests" of the Illinois Association of County Superintendents met in conference at the University of Illinois, in order to decide upon a testing program for rural schools. A group of three tests, including one for the measurement of general intelligence, one for measuring ability to read silently, and one for measuring abilities in the field of arithmetic, were recommended to the committee by the Bureau of Educational Research. Although the school year was far advanced it was decided to incorporate in this battery of tests, the achievement quotient which had been conceived in another connection. ✓

It is obvious that educational research of the type represented by the bulletin is possible only through the cooperation of a large number of persons. The writer is glad to acknowledge the cooperation of, *first*, the city superintendents and teachers who gave the preliminary form of the tests in May, 1920, *second*, the county superintendents and others who cooperated in the giving of the tests in October, 1920, and *third*, those city superintendents and teachers who gave portions of the Examination in the spring of 1921 in order to determine the reliability of the tests. Acknowledgement is made in the body of the bulletin for special contributions.

WALTER S. MONROE, Director.

THE ILLINOIS EXAMINATION.

INTRODUCTION

The Illinois Examination, published in the form of a sixteen page booklet, is the name given to a battery of tests: The Illinois General Intelligence Scale, Monroe's Standardized Silent Reading Tests, Revised, and Monroe's General Survey Scale in Arithmetic.* However, the significant characteristic of the Illinois Examination is not that of merely putting these three tests together in one booklet. It is rather the way in which the measures of achievement are combined with the measure of intelligence to secure the Achievement Quotient (A. Q.) The usual procedure of interpreting measures of achievement by reference to grade norms provides for no consideration of the general intelligence of the pupil; all pupils, bright, average, and dull are judged with reference to the same norms. The plan by which the measures of achievement and general intelligence are combined in the Illinois Examination was conceived in an effort to provide a procedure by which a pupil's general intelligence or his capacity to learn would be considered in interpreting his achievement.

Briefly, the plan consists of establishing for the achievement tests mental age norms which are used to supplement the usual grade norms. For each half year of mental age, as shown by the general intelligence scale used, the median achievement was determined. These medians are the mental age norms, which are used as a basis for translating the point scores into achievement ages. In arriving at the mental age norms, all pupils of a given mental age in grades III to VIII inclusive were grouped together without regard to the grade in which they were classified. Provision is made for comparing a pupil's achievement score, (when expressed as an achievement age) with the norm corresponding to his mental age by dividing his achievement age by the standard score for his mental age.** This quotient is called the Achievement Quotient. The plan involves certain assumptions and approximations. These, as well as the validity and the reliability of the several tests incorporated in the Illinois Examination, will be considered in detail on the following pages.

In the construction of the tests crude methods were frequently employed in preference to more refined ones because of their greater simplicity and the saving of time thus secured. For example, this was done in the construction of the duplicate forms of the tests in silent reading and arithmetic and in deriving the basis for translating the point scores into age scores. In the latter case a more refined procedure involving the use

*These tests are published by the Public School Publishing Co., Bloomington, Illinois.

**This "standard score" is numerically the same as a pupil's mental age.

of the regression equation might have been used. The results obtained by the use of the crude methods have been verified by a critical study of the test.

I. FACTS OF TITLE.

In so far as the Illinois General Intelligence Scale represents originality, B. R. Buckingham is primarily responsible for it. Walter S. Monroe, the author of this bulletin, devised the two achievement tests, and also contributed the plan of combining measures of achievement with measures of intelligence. An experimental edition of the Illinois Examination, including the scales for general intelligence and silent reading, was published in April, 1920. The complete examination together with the Teachers Handbook and class record sheet was first published during the summer of 1920 and was made available for distribution in August. The second form of the Illinois Examination was devised during the school-year of 1920-21 and was made available for use early in 1921. The Examination consists of two parts. Illinois Examination I is designed for grades III, IV and V. Illinois Examination II is designed for grades VI, VII, and VIII. The Illinois General Intelligence Scale is the same for all grades. The scales for reading and for arithmetic are different in the two examinations. There are two forms of the examination. A third form of the scales for silent reading and arithmetic has been prepared.

II. NATURE OF PUPIL'S PERFORMANCE.

1. **Illinois General Intelligence Scale.** In the Illinois General Intelligence Scale the pupil is required to make a mark with his pencil or to write figures. No other writing is required of him. The marks consist either of a line drawn under a word or of a line crossing out a word or number.

The scale consists of seven sub-tests which are illustrated by a few samples from each.

Test No. 1—ANALOGIES

1	eat—bread:: drink— water iron lead stones.....	1
2	finger—hand::toe— box foot doll coat.....	1
3	shoe—foot::hat— kitten head knife penny.....	3
4	dress—women::feathers— bird neck feet bill	4
5	dog—puppy::cat— kitten dog tiger house.....	5

The pupil is asked to draw a line under the word in heavy type related to the third word as the second is related to the first.

Test No. 2—ARITHMETIC PROBLEMS

- 1 If one boy has 10 fingers, how many fingers have six boys?..... Answer ()
- 2 There are 15 children in our class. 5 of them are boys. How many are girls?..... Answer ()
- 3 We learn 2 words a day in our class. How many do we learn in 8 days? Answer ()
- 4 Jack is 42 inches tall and Fred is 5 inches taller. How tall is Fred? Answer ()
- 5 Mr. Gray sold ten bags of flour last Saturday at 2 dollars a bag.
How many dollars did he get for the flour? Answer ()

The pupil is required to write the answer within the parentheses.

Test No. 3—SENTENCE VOCABULARY

- 1 A gown is a string animal dress plant.
- 2 Haste is hurry red little sweet.
- 3 To tap is to run fall knock smile.
- 4 A dungeon is open bright heavy dark.
- 5 Majesty refers to dresses kings countries climates.

The pupil is required to draw a line under that one of the last four words which makes a true sentence.

Test No. 4—SUBSTITUTION

7
1

7
2

7
3

7
4

7
5

7	
7	
7	
7	
7	

7	7		
7	7		
7	7		
7	7		
7	7		

7	7	7			
7	7	7			
7	7	7			
7	7	7			
7	7	7			

The pupil is to write in squares after the symbols the numbers to which they correspond as given by the key at the top.

Test No. 5—VERBAL INGENUITY

- 1 the cat at see.
- 2 boy was sky the sick.
- 3 Bread sweep will the kitchen I.
- 4 are going yesterday to-morrow we.
- 5 me mine give my straw hat.

The pupil is to cross out the superfluous word. The remaining words can then be made into a true sentence.

Test No. 6—ARITHMETICAL INGENUITY

- (1) 1 2 3 9 4 5
- (2) 2 4 6 7 8
- (3) 9 8 7 6 5 2
- (4) 11 10 8 6 4 2
- (5) 5 7 10 15 20 25

The pupil is to cross out the number that does not fit in the group.

Test No. 7—SYNONYM-ANTONYM

- | | | | |
|---|----------------------|---------------|---|
| 1 | high—low..... | same—opposite | 1 |
| 2 | go—leave | same—opposite | 2 |
| 3 | large—great | same—opposite | 3 |
| 4 | bitter—sweet..... | same—opposite | 4 |
| 5 | begin—commence | same—opposite | 5 |

The pupil is required to indicate whether the first two words mean the same or the opposite by drawing a line under the word “same” or “opposite.”

These sub-tests were selected after analyzing a number of existing general intelligence scales with reference to sub-tests and after examining much that has been written on the topic of group intelligence scales. The opinion of certain experts in this field was also obtained. Test No. 1, Analogies, was made from exercises taken from the Army Scale Alpha. Test No. 2, Arithmetic Problems, was compiled from a collection of problems which Professor Buckingham had evaluated with reference to difficulty. Test No. 3, Sentence Vocabulary, is an abridgment of the Holley Sentence Vocabulary Scale. In general this abridgment was secured by taking every third exercise of the original test, beginning with the first. Test No. 4, Substitution, was contributed by Professor E. H. Cameron of the College of Education, University of Illinois. Test No. 5, Verbal Ingenuity, and Test No. 6, Arithmetical Ingenuity, were taken from the Pressey Cross Out Tests.* Test No. 7, Synonym-Antonym, consists of exercises selected largely from the Army Scale Alpha.

These seven sub-tests together with two others were given to a limited number of pupils before the preliminary edition was printed. On the basis of these data two of the sub-tests were eliminated and minor changes made in the seven retained. In general the exercises in each of the sub-tests are arranged in order of gradually increasing difficulty.

*Acknowledgment is hereby made to Professor S. L. Pressey, of Ohio State University, for permission to use these tests.

Form 2 of the sub-tests was devised in either of two ways. The first method was to make use of "standardized exercises," either in the files of the Bureau of Educational Research, or available in published reports.* The second method involved the modification of the exercises of the first form in such a way as to change their identity with the least probable change in difficulty.

The first method was used in the case of Test 1, Analogies; Test 2, Arithmetic Problems; Test 3, Sentence Vocabulary; and Test 7, Synonym-Antonym. For Test 1, exercises published by Professor R. Pintner in the *Journal of Applied Psychology*, June-September, 1920, were used. For Test 2, exercises were taken from Professor Buckingham's own list of such material. The exercises for Test 3 were taken from a list of such exercises evaluated by Dr. C. E. Holley when he was engaged in the derivation of his Sentence Vocabulary Scale. The exercises for Test 7 were taken from the Army Scale Alpha as was done in the case of Form 1.

Form 2 of Test 4, Substitution, was constructed by using the same characters as in Form 1, but changing their numerical equivalents. The second form of Test 5, Verbal Ingenuity, and of Test 6, Arithmetical Ingenuity, are merely modifications of Form 1. Several modifications of each of the exercises were given to pupils and that one selected which proved to be most nearly of the same difficulty as the corresponding exercise of Form 1.

The pupil is made acquainted with the nature of the exercises in each sub-test by means of a few sample exercises in addition to a verbal explanation. With the exception of Test 2, Arithmetic Problems, three or four sample exercises are given for each sub-test. The time allowance for each of the sub-tests is given below. In each case it is intended that no pupils, even the brightest, will finish all of the exercises within the time allowed.

Test 1, Analogies,	2 minutes
Test 2, Arithmetic Problems,	3 "
Test 3, Sentence Vocabulary,	2 "
Test 4, Substitution,	3 "
Test 5, Verbal Ingenuity,	3 "
Test 6, Arithmetical Ingenuity,	2 "
Test 7, Synonym-Antonym,	1 "
Total	16 minutes

Equivalence of duplicate forms. In order to determine the degree of equivalence of the two forms of the Illinois General Intelligence Scale, copies of the two forms were arranged in alternate order and distributed to

*This standardization refers to the determination of the difficulty of the exercises upon the basis of the percent of pupils who were able to do them correctly.

pupils as they happened to be seated. This was done in a number of school systems. The median and average scores secured from the two forms in this way are given in Table I. The median and average scores are expressed in terms of points. The differences for both the median and the average show that the forms are approximately equal. These differences vary slightly from grade to grade. The greatest difference exists in the fourth grade. Except in two instances the differences are positive which shows that the second form yields slightly larger scores than the first. This means that the second form is slightly easier than the first. When the six grades are combined the difference between the medians is 0.8. The difference of the averages is 2.3. Since, as we show later, 10.0 points of a pupil's score are equivalent to one year of mental age, we may say that in general the non-equivalence of the two forms of this scale is probably not more than two months.

TABLE I. DEGREE OF EQUIVALENCE OF FORMS 1 AND 2 OF ILLINOIS GENERAL INTELLIGENCE SCALE

Grade	No. of Pupils		Median Score		Difference	Average Score		Difference
	Form 1	Form 2	Form 1	Form 2	Form 2-Form 1	Form 1	Form 2	Form 2-Form 1
III	331	325	32.3	32.7	0.4	33.1	33.7	0.6
IV	298	295	46.0	50.5	4.5	47.3	51.5	4.2
V	336	334	56.9	56.7	-.2	58.0	59.4	1.4
VI	300	288	73.7	76.2	2.5	75.1	76.0	0.9
VII	289	279	81.9	84.1	2.2	82.6	83.8	1.2
VIII	240	253	102.1	101.3	-.8	101.3	102.7	1.4
III-VIII	1794	1774	62.5	63.3	0.8	63.8	66.1	2.3

2. **Monroe's Standardized Silent Reading Tests, Revised.** In these tests the pupil is asked to read a series of exercises which have no connection with each other. In each exercise the pupil is required to read a paragraph and to answer a question based upon it. The answer is to be given by drawing a line under a word or by indicating it in some other way. No writing is required. In most exercises the pupil has to select one out of five words in making his response; in a few there are only four words. The nature of the exercises may be illustrated by the following. The first two are taken from Test I and the others from Test II:

It was a rainy, dark, dismal day. The children had not been allowed to go out to play all day. Their lessons were poor and the teacher cross. It was late in the afternoon.

Draw a line under the word that tells how the children felt.

active smiling happy cross good

"The golden rod is yellow,
The corn is turning brown,
The trees in apple orchards
With fruit are bending down."

Draw a line under the season of the year you think is pictured in this stanza.

autumn spring winter summer

It was cold, bleak, biting weather; foggy withal; and he could hear the people in the court outside go wheezing up and down, beating their hands upon their breasts and stamping their feet upon the pavement-stones to warm them.

What kind of picture does this paragraph describe?

comfortable luxurious cheerless pleasant exciting

The caravan, stretched out upon the desert, was very picturesque; in motion, however, it was like a lazy serpent. By and by its stubborn dragging became intolerably irksome to Balthasar, patient as he was.

Place a line under the word which tells in what respect the caravan resembled a serpent.

temper color length motion size

In front the purple mountains were rising up, a distant wall. Cool snow gleamed upon the summits. Our horses suffered bitterly for water. Five hours we had ridden through all that arid waste without a pause.

What kind of a country had these people been riding through?

mountainous swampy desert forest valley

The exercises for Form 1 were taken with some modifications from the original edition of Monroe's Standardized Silent Reading Tests. In selecting the exercises, and in making the modifications, an effort was made to have all the exercises approximately the same length. Those for test II are slightly longer than the ones included in Test I. The exercises are not absolutely equal with respect to difficulty. They are arranged so that in general there is a slight increase in difficulty from exercise to exercise but in no sense can they be considered a difficulty scale. To secure absolute uniformity with respect to difficulty would have required the expenditure of a prohibitive amount of labor. Even if this were not true it is believed to be desirable to have a moderate range of difficulty in view of the fact that the exercises are to be given to pupils in a sequence of three successive school grades.

The nature of the exercises is explained to the pupils by means of three fore-exercises. In addition they are given certain verbal explanations. Four minutes are allowed for the test in all grades. This time allowance was intended to be such that practically no pupils would complete the tests; but it has been found that this is not always the case.

A few of the exercises are based on poetry although the majority are based on prose paragraphs. In light of a recent investigation* it appears likely that the reading of poetry, even for the purpose of answering a question, is not the same activity as the reading of prose for the same purpose. To the extent that this is true the test is not consistent, and this constitutes one of its limitations. In a number of other respects a high degree of uniformity has been secured. In every case the pupil's response is the same. The exercises are approximately equal in length and the questions asked appear to call for much the same type of reading.

Equivalence of duplicate forms. The three forms of each of the silent reading tests were arranged in alternate order and given in this order to pupils as they were seated in a number of cities. Table II gives the median and average comprehension scores for each of the three forms.** With few exceptions these differences are less than one. A few are negative but most are positive. In the case of Test I, when the scores from grades III to V inclusive are combined, the three forms are shown to be very nearly equivalent both by the difference of the medians and by the difference of the averages. In the case of Test II, which is designed for grades VI, VII, and VIII, Forms 1 and 2 are shown to be exactly equivalent by the medians and approximately so by the averages. Form 3, however, appears to be slightly easier, thus yielding slightly larger scores than either Form 1 or Form 2. The difference approaches a year of achievement in comprehension.***

In Table III, corresponding data are given for the rate scores. The**** differences here are larger. This is partly due to the fact that rate scores are expressed in terms of a much smaller unit than comprehension scores, but in this test the reading rate appears to be more erratic than comprehension. Although some of the differences for Test I are relatively small, corrections should be made when comparisons are made between scores obtained from the different forms.

3. **Monroe's General Survey Scale in Arithmetic.** Scale I, designed for grades III, IV, and V, consists of eight sub-tests. Scale II, designed for grades VI, VII, and VIII, consists of seven sub-tests. In each sub-test, except No. 7, Scale II, the pupil is asked to do arithmetical examples. In Test 7, he is asked to insert the decimal point in quotients.

*Pressey, L. W. and Pressey, S. L. "A critical study of the concept of silent reading ability." *Journal of Educational Psychology*, 12:25-31, January, 1921.

**The unit in terms of which scores are expressed is relatively large. It approaches one year of achievement in comprehension of silent reading.

***Since Form 3 is not incorporated in the Illinois Examination, its non-equivalence with the other two forms does not constitute a limitation of this battery of tests.

****Except for the lower achievement ages, 7 or 8 units are equivalent to one year of achievement in rate of silent reading.

TABLE II. DEGREE OF EQUIVALENCE OF FORMS I AND 2 OF MONROE'S STANDARDIZED
SILENT READING TESTS, REVISED (COMPREHENSION)

Grade	Number of Pupils			Median Score			Difference in Forms			Average Score			Difference in Forms		
	Form 1	Form 2	Form 3	Form 1	Form 2	Form 3	2-1	3-1	3-2	Form 1	Form 2	Form 3	2-1	3-1	3-2
III IV V	323	299	321	5.2	4.9	5.1	-0.3	-1.1	0.2	5.8	5.3	5.4	-0.5	-0.4	0.1
	286	303	299	7.4	7.4	7.3	0.0	-1.1	-1.1	7.5	7.8	7.8	0.3	0.3	0.0
	340	302	325	9.2	9.5	10.1	0.3	0.9	0.6	9.6	10.0	10.1	0.4	0.5	0.1
III-V	949	904	945	7.6	7.5	7.4	-0.1	-0.2	-0.1	7.7	7.7	7.8	0.0	0.1	0.1
VI VII VIII	269	272	275	11.5	11.9	11.7	0.4	0.2	-0.2	11.4	12.0	12.2	0.6	0.8	0.2
	279	265	277	13.1	12.6	13.4	-0.5	0.3	0.8	13.2	12.9	13.5	-0.3	0.3	0.6
	249	267	255	14.5	15.0	16.5	0.5	2.0	1.5	14.5	15.1	16.4	0.6	1.9	1.3
VI-VIII	797	804	807	13.0	13.0	13.6	0.0	0.6	0.6	13.0	13.3	14.0	0.3	1.0	0.7

TABLE III. DEGREE OF EQUIVALENCE OF FORMS 1 AND 2 OF MONROE'S STANDARDIZED SILENT READING TESTS,
REVISED (RATE)

Grade	Number of Pupils			Median Score			Difference in Forms			Average Score			Difference in Forms		
	Form 1	Form 2	Form 3	Form 1	Form 2	Form 3	2-1	3-1	3-2	Form 1	Form 2	Form 3	2-1	3-1	3-2
III IV V	323	299	321	86.9	82.3	85.6	-4.6	-1.3	3.3	90.2	83.8	93.1	-6.4	2.9	9.3
	286	303	299	113.7	109.0	110.1	-4.7	-3.6	1.1	113.6	118.2	117.2	4.6	3.6	-1.0
	340	302	325	138.7	129.1	143.6	-9.6	4.9	14.5	133.4	134.8	141.0	1.4	7.6	6.2
III-V	949	904	945	116.9	104.7	110.0	-12.2	-6.9	5.3	112.7	112.3	117.2	-4	4.5	4.9
VI VII VIII	269	272	275	155.2	160.7	168.8	5.5	13.6	8.1	153.9	161.5	170.5	7.6	16.6	9.0
	279	265	277	166.5	167.4	181.0	0.9	14.5	13.6	172.3	169.0	183.0	-3.3	10.7	14.0
	249	267	255	187.2	189.8	223.2	2.6	36.0	33.4	187.4	196.8	225.5	9.4	38.1	28.7
VI-VIII	797	804	807	165.7	173.0	183.5	7.3	17.8	10.5	170.8	175.6	192.2	4.8	21.4	16.6

In selecting the sub-tests for each scale, an effort was made to include examples of the types most appropriate for the pupils to whom they would be given. Tests 1, 2, 3, and 4 of Scale I are on the fundamental combinations, or tables, one test being devoted to each operation. These tests are similar to the corresponding tests of the Courtis Standard Research Tests, Series A, and of the Cleveland Survey Tests. Test 5 calls for single column addition of five figures. Test 6 consists of subtraction examples, in which the subtrahend is a single figure. Test 7, multiplication, consists of multiplication examples in which the multiplier is a single figure. In Test 8, division, the divisor is a single figure.

The sub-tests of Scale II are represented by the following samples:

Test No. 1—ADDITION

7862	6809	8941	5917	6772	7864	1249
5013	7623	7910	4814	6028	7883	8975
1761	5299	9845	9007	6535	8240	9005
5872	6601	8522	6975	2340	9869	1573
3739	3496	1046	1227	2319	6794	3203

Test No. 2—MULTIPLICATION

4857	5718	6942	4065
36	92	58	47

Test No. 3—DIVISION

41)574	79)36893	32)384	58)27608
--------	----------	--------	----------

Test No. 4—SUBTRACTION

739	1852	975	1087	516	962
367	948	906	821	239	325

Test No. 5—ADDITION AND SUBTRACTION OF FRACTIONS

$$\frac{1}{6} + \frac{1}{3} =$$

$$\frac{3}{4} - \frac{2}{5} =$$

$$\frac{1}{6} + \frac{3}{5} =$$

Test No. 6—MULTIPLICATION AND DIVISION OF FRACTIONS

$$\frac{2}{3} \times \frac{3}{4} =$$

$$\frac{4}{7} \div \frac{2}{3} =$$

$$\frac{5}{12} \times \frac{3}{5} =$$

Test No. 7—DECIMAL FRACTIONS

.03)16.2	Ans.: 54	.07)1.82	Ans.: 26	.05).415	Ans.: 83
.06)7.44	Ans.: 124	.08).952	Ans.: 119	.04)87.6	Ans.: 219
.02).144	Ans.: 72	.08)40.8	Ans.: 51	.09)3.42	Ans.: 38

Several of the sub-tests were taken bodily from Monroe's Diagnostic Tests in Arithmetic. Others were constructed similar to well known tests. In constructing the duplicate forms of these tests the figures of the examples were rearranged or changed, so that examples identical in gross structure but leading to different answers were obtained. In the case of Scale I, Tests 1, 2, 3, and 4 were constructed by securing at random numbers of the same general magnitude as those used in Form 1. In doing this, especially in Tests 1 and 3, many of the examples were repeated either in their identical form or with the position of the two numbers reversed. Test 5 was constructed by rearranging the same numbers actually used in Test 5 of Form 1. This rearrangement was largely, although not entirely so, a process of shifting the position of the example and reversing the order of numbers therein. In Test 6, two methods were used, either the subtrahend and minuend were grouped together differently or the figures in the minuend were reversed. In Test 7, the figures in the multiplicand were rearranged and the resultant number grouped with another multiplier. In Test 8, the same divisors were kept and the dividends either slightly increased or decreased so as to leave the quotients still integral numbers.

In Test 1 of Scale II, the columns of figures in each example were shifted with occasional changes to prevent a zero from coming first in any number. In Test 2, the figures of the multiplicand were rearranged and the resultant number grouped with another multiplier. In Test 3, either the dividends were slightly increased or decreased so as to be evenly divisible by the same divisor, or the dividend or divisor was multiplied or divided by two. In Test 4, either the figures of the minuend were arranged differently or new minuends were constructed by grouping together part of the figures from two minuends to form one, the remaining figures to form another. The subtrahends were either changed by a different arrangement of figures or by adding or subtracting a small number, such as ten or twenty, or they were not changed at all. Tests 5 and 6 were constructed by a random selection of fractions of the same general magnitude as those of the same test in Form 1. In Test 7, the position of the decimal point in either divisor or dividend or in both was shifted, and the position of the examples was also changed.

In almost all of the tests an occasional change, not covered by the statements above, was made, either because of the necessity of avoiding impossible combinations or because the result by following too closely the procedure laid down seemed undesirable.

Equivalence of Duplicate Forms. A study was made of the equivalence of the duplicate forms of Monroe's General Survey Scale in Arithme-

tic in the same way that the equivalence of the duplicate forms of tests for general intelligence and silent reading was investigated. The results are given in Table IV. Since Forms 2 and 3 were constructed from Form 1 by a rather mechanical procedure one might expect to find a higher degree of equivalence for the arithmetic scale than for the others. This, especially in the upper grades, is not true. In the case of Scale I, however, the differences are relatively small, particularly as shown by the averages.* In the case of Scale II the differences between Form 1 and Form 2 are small. Hence, these two forms may be considered as being very approximately equivalent. Form 3 appears to be considerably easier than the other two.

III. DESCRIPTION OF PUPIL'S PERFORMANCE.

Point Scores. The performance of a pupil upon each of the sub-tests of the Illinois General Intelligence Scale is described in terms of a point score. With two exceptions this point score is the number of exercises which he has done correctly in the time allowed. In order to give appropriate weight to Test No. 4, Substitution, the number of figures written correctly is divided by four. Since the possible score on this test is 150, as compared with 16 to 40 in the cases of the other tests, failure to weight the score from this test would give it an exceedingly high degree of potency in determining the pupil's total score. The rule to divide this score by 4 is purely arbitrary. In Test No. 7, Synonym-Antonym, in order to discount for the correct answers which might be obtained by merely guessing, the pupil's score is given as the number of exercises right minus the number wrong; exercises skipped are counted as wrong; exercises not reached when time is called are not counted. The pupil's total point score is the sum of his scores on the seven sub-tests.

In the case of Monroe's Standardized Silent Reading Tests, Revised, the pupil receives two scores. His comprehension score is the number of exercises which he answers correctly. His rate score is the number of words which he reads per minute or the total number of words read divided by 4. In order to obtain this rate the pupil is asked to mark the line which he is reading when the signal to stop is given. The cumulative totals of the words are printed in the left hand margin so that the total number of words read is easily obtained.

The number of examples right is taken as the pupil's point score on each of the sub-tests of Monroe's General Survey Scale in Arithmetic. These sub-tests yield scores which differ widely in magnitude. Arbitrary

*The equivalence of one year of achievement is 8 units of the point score.

TABLE IV. DEGREE OF EQUIVALENCE OF FORMS 1 AND 2 OF MONROE'S GENERAL SURVEY SCALE IN
ARITHMETIC

Grade	Number of Pupils			Median Score			Difference in Forms			Average Score			Difference in Forms		
	Form 1	Form 2	Form 3	Form 1	Form 2	Form 3	2-1	3-1	3-2	Form 1	Form 2	Form 3	2-1	3-1	3-2
III IV V	225	222	227	19.6	20.6	20.5	1.0	0.9	-.1	20.9	21.2	22.1	0.3	1.2	0.9
	205	207	188	33.1	37.3	33.9	4.2	0.8	-3.4	38.1	40.1 ^{1/2}	38.5	2.0	0.4	-1.6
	219	223	219	48.4	50.8	47.2	2.4	-1.2	-3.6	50.5	51.0	49.4	0.5	-1.1	-1.6
III-V	649	652	634	32.8	34.8	33.7	2.0	0.9	-1.1	33.1	34.0	33.0	0.9	-.1	-1.0
VI VII VIII	222	217	210	53.6	56.5	62.6	2.9	9.0	6.1	55.5	58.8	66.4	3.3	10.9	7.6
	204	203	201	61.5	59.9	68.9	-1.6	7.4	9.0	63.9	62.3	70.1	-1.6	6.2	7.8
	192	202	198	78.0	80.0	75.7	2.0	-2.3	-4.3	82.1	81.6	76.0	-.5	-6.1	-5.6
VI-VIII	618	622	609	64.1	64.2	68.3	0.1	4.2	4.1	66.5	67.0	70.6	0.5	4.1	3.6

rules were adopted for weighting these scores so that approximately equal weight would be given to each test. (For details see the tests.) The weighted sum of the scores on the several sub-tests is the pupil's point score.

Correction of scores derived from the scales designed for grades VI, VII, and VIII. In the cases of both silent reading and arithmetic, the tests for grades III, IV, and V are entirely different from those given in grades VI, VII, and VIII. This makes the scores from the two sequences of grades incomparable. They have a different zero point and it is not unreasonable to expect that they would be expressed in terms of a different unit. It is relatively easy to estimate the approximate difference between the zero points. This difference can be used as a correction to be added to the point scores obtained from the tests for the upper sequence of grades. This was done and it appears that the differences in the units are not sufficiently large to introduce serious inaccuracies when the corrected scores are considered comparable.

The method of estimating the difference between the zero points may be illustrated by the case of silent reading. The medians of the point scores without correction were calculated for each grade. These were then represented graphically as shown in Figure 1. The curve of progress for grades III, IV, and V was then extended so that the extension would parallel the curve of progress for grades VI, VII, and VIII. This extension together with the progress curve for the lower sequence of grades forms a progress curve for grades III to VIII inclusive. The distance between the extension and the original curve is assumed to represent the difference in the zero points of the two tests.

The estimated corrections based upon the scores derived from the preliminary tests in May, 1920, for Monroe's Standardized Silent Reading Test II, Revised, were rate 17 and comprehension 4. The arithmetic tests were not given at this time but, on the basis of scores derived from similar tests, a correction of 16 was estimated as the proper amount to add to the point scores derived from Scale II. The scores derived from giving the Illinois Examination to fifty thousand children during the fall of 1920 indicated that these estimates were incorrect. The revised estimates are: reading rate 29, reading comprehension 5, and arithmetic 22. When these corrections are added to the point scores derived from the corresponding tests in Examination II, the scores will, in general, be approximately comparable to the corresponding scores derived from Examination I.

2. Derived scores. The point scores yielded by the scales of the Illinois Examination are translated into age scores. From these quotients

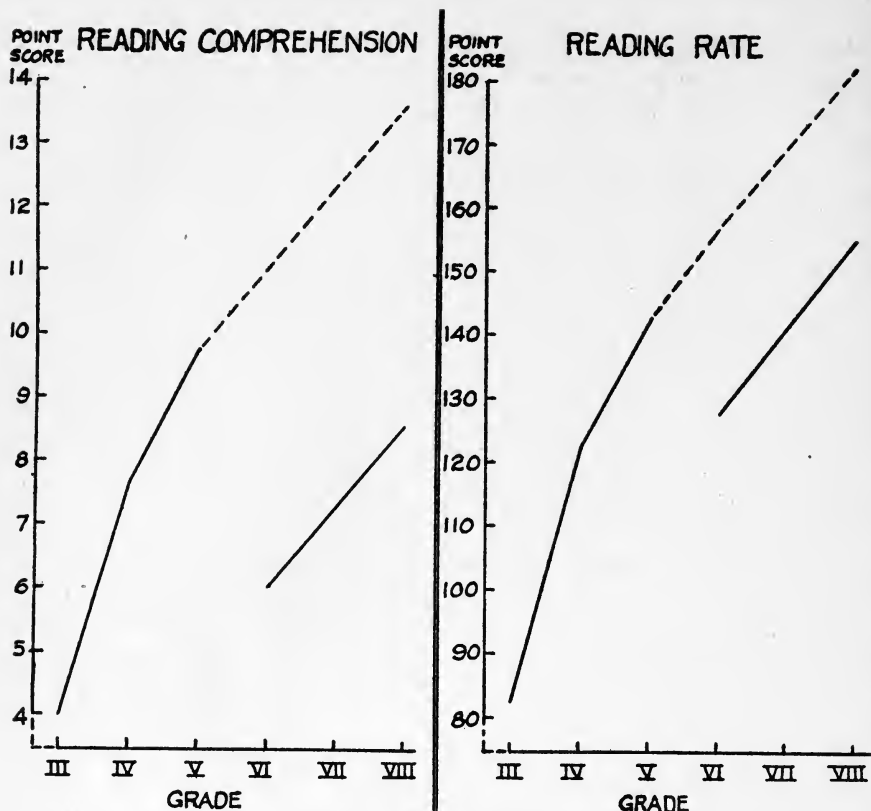


FIGURE 1. SHOWING METHOD OF ESTIMATING CORRECTIONS TO BE ADDED TO SCORES OF SCALE II.

are calculated. The portion of the Illinois Examination devoted to the measurement of general intelligence and ability in silent reading was given to over seventy-five hundred pupils in May, 1920, for the purpose of standardization and of determination of the basis for calculating certain derived scores. When the complete examination was used in October, 1920, the scores obtained indicated that the results secured in May involved a sufficiently large error to make necessary a re-determination of the grade norms and of the basis for calculating the derived scores. Unfortunately, the data collected from the October testing were in such form that the determination of the basis for calculating the derived scores could not be worked out by direct methods. It was, therefore, necessary to use estimates instead of actual determinations, but these estimates have been verified whenever possible and appear to be approximately correct.

Data, in the form required for the direct calculation of the basis for obtaining the derived scores, are available from the public schools o.

Decatur, Illinois, and from eight elementary schools of the city of Chicago. These data, particularly those from Decatur, do not appear to be entirely representative. However, it has seemed wise to describe the method in terms of the scores from these two cities. The estimates will be found to agree, in some instances, with the data collected in these two cities, but in others agreement will be lacking.

The chronological age norms for the Illinois General Intelligence Scale are used as a basis for translating the point scores into mental age scores. In determining the age norms the distribution of point scores was secured for each half year of chronological age. For example, the point scores of all pupils whose chronological ages fell between nine years, and nine years and six months, were brought together in a distribution. The median of this distribution was taken as the age norm for this chronological age group. In Table V, the median point scores for each of the age groups are given for both Decatur and Chicago. In the second column of the table the average chronological age of the group is given. It will be noted that except for the extreme ages the median point scores tend to become larger as the pupils become older.

This scale was given only to the pupils in grades III to VIII inclusive. If a pupil enters school at six years we may expect to find him in the third grade at eight years of age. Of course, if the pupil fails of promotion, he will not reach the third grade until he is older. Pupils younger than eight years attaining the third grade either have entered school before they were six or have skipped a grade. In this case the pupils are probably bright. The same thing is true to a somewhat less degree for the pupils whose ages fall in the 8.0-8.4 interval, and even in the interval from 8.5-8.9 there is some selection of the brighter pupils. All pupils eight years of age have not advanced to the third grade. Some of them have failed of promotion and are found in the first or second grade. In general, they are the duller pupils. Consequently, we may say that the age groups below nine years are selected so that they do not include all pupils whose ages fall within the groups. The less capable have been left out. In Chicago, the Illinois General Intelligence Scale was not given to III-B pupils. For this reason, the selection extends above nine years.

A pupil who entered school at six years of age and was promoted each year would be found at fourteen in the ninth grade or first year of the high school. Furthermore, when pupils reach the age of fourteen, the compulsory attendance law does not apply with the same force as in the case of younger pupils. For these reasons the age-groups at the top of the table are selected. In general, they include pupils who have failed of promotion one or more times because they were not able to do the work

of the school satisfactorily. Hence, we should expect to find the median point scores for such pupils less than those for younger groups of unselected pupils.

TABLE V. MEDIAN POINT SCORES OF CHRONOLOGICAL AGE GROUPS FOR DECATUR AND FOR EIGHT ELEMENTARY SCHOOLS IN CHICAGO

Chronological Age Interval	Average Chronological Age	Decatur		Chicago	
		No. of Pupils	Median Point Score	No. of Pupils	Median Point Score
17.5-17.9	17.75	1	40.0		
17.0-17.4	17.25				
16.5-16.9	16.75	2	54.5	5	57.5
16.0-16.4	16.25	9	70.0	11	61.7
15.5-15.9	15.75	26	64.5	37	67.9
15.0-15.4	15.25	61	64.4	78	75.3
14.5-14.9	14.75	124	71.4	130	72.4
14.0-14.4	14.25	191	68.9	284	70.0
13.5-13.9	13.75	244	73.5	366	67.6
13.0-13.4	13.25	297	70.2	408	72.2
12.5-12.9	12.75	306	67.2	326	67.1
12.0-12.4	12.25	304	63.4	354	63.1
11.5-11.9	11.75	308	52.8	294	57.8
11.0-11.4	11.25	298	54.8	328	55.2
10.5-10.9	10.75	316	43.8	244	46.9
10.0-10.4	10.25	290	40.9	313	48.0
9.5-9.9	9.75	297	35.9	205	44.4
9.0-9.4	9.25	285	30.1	168	43.2
8.5-8.9	8.75	216	26.2	78	40.0
8.0-8.4	8.25	163	20.0	24	38.3
7.5-7.9	7.75	33	27.5	3	42.5
7.0-7.4	7.25	14	23.0		
6.5-6.9	6.75	2	20.0		

In Figure 2, the median point scores given in Table V are represented graphically. The Decatur scores are represented by small crosses and the Chicago scores by small circles. With the exception of the scores for the older pupils there is a suggestion of a straight line relationship between the point scores and chronological age. Other considerations, some of which will be mentioned later, led to the adoption of a straight line as representing the relation between the median point scores and the chronological ages. This line has been drawn in the figure. It is considered to extend indefinitely upward. It is extended downward until it cuts the vertical axis. This straight-line relationship permits the formulation of a very simple rule for translating point scores into mental ages.

cates that beginning about 14 years there is a slowing down of development and that at about the age of 16 or 18 it ceases altogether* After this age, mental changes consist primarily in increasing the scope and skillfulness of the application of intelligence which one already possesses. An extension of the line of relationship which might be judged to conform with the actual median point scores for the higher age groups (assuming that all pupils belonging to these age groups were tested) would be curved downward and would tend to become horizontal or nearly so. This procedure, which is the one followed in the Binet Scale and in a number of others, would not give us a method for translating the point scores made by the bright children in the upper grades into the corresponding ages. For example, a few pupils have made point scores as high as 185 on the Illinois General Intelligence Scale. On the basis of the facts at hand it is impossible to conceive that the median point score for any age group could possibly equal this. Therefore, if such point scores are to be translated into mental ages it is necessary to devise another plan. The plan adopted was to extend the straight line of relationship upward until maximum point scores were reached. This plan has the merit of simplicity.

If the downward extension of the line of relationship is examined it will be found that the point score which is equivalent to zero mental age is -55. This means that the absolute zero of the Illinois General Intelligence Scale is -55 points. Therefore, in translating point scores into corresponding mental ages, it is necessary to correct this by adding 55 to the point score. When the slope of the line of relationship is examined it will be noted that the rise is approximately 10 points for one year of chronological age, which means that 10 points on the intelligence scale is considered equivalent to one year of mental age. Hence, we are able to state the rule, "Add 55 to the total point score and divide by 10. The quotient is the corresponding mental age." This rule is not absolutely accurate but it does not appear to involve errors greater than the errors of measurement which will be discussed in a later section.

The validity of the rule for determining the mental ages was verified by comparing the results obtained by using it with the mental ages as determined by the Stanford Revision of the Binet Scale for Measuring Intelligence. Table VI gives the determinations of mental age by the two scales for 201 pupils.** Because of the small number of pupils in some

*Children probably differ with reference to the development of their general intelligence. This slowing down probably begins at different ages in the cases of different pupils. Thus this statement should be interpreted as representing the average or typical development.

**A portion of these data was furnished the writer by Professor J. C. DeVoss, of the Kansas State Normal School, Emporia, Kansas. The remainder was contributed by Superintendent L. W. Keeler, Michigan City, Indiana.

groups it is to be expected that some of the differences in mental age would be relatively large. An examination of the differences given in the last column of the table reveals that only three out of twenty are greater than one year. Many of the differences are very small. This table indicates that the rule adopted for translating point scores into mental ages gives results agreeing, as closely as could be expected in view of the errors of measurement which occur in using both scales, with the results obtained by using the Stanford Revision of the Binet Scale for Measuring Intelligence.

TABLE VI. COMPARISON OF MENTAL AGES OF 201 PUPILS BY BINET SCALE AND BY ILLINOIS GENERAL INTELLIGENCE SCALE

Number of Pupils	Binet Scale		Illinois Scale		Difference in Mental Ages
	Mental Age Interval	Mental Age Average	Med. Point Score	Corresp'd'g Mental Age	
2	17.5-17.9	17.75	125.0	18.00	-.25
	17.0-17.4	17.25			
2	16.5-16.9	16.75	115.0	17.00	-.25
2	16.0-16.4	16.25	130.0	18.50	-2.25
8	15.5-15.9	15.75	90.0	14.50	1.25
6	15.0-15.4	15.25	105.0	16.00	-.75
5	14.5-14.9	14.75	75.0	13.00	1.75
9	14.0-14.4	14.25	85.0	14.00	.25
13	13.5-13.9	13.75	92.1	14.71	-.96
19	13.0-13.4	13.25	78.8	13.38	-.13
11	12.5-12.9	12.75	71.7	12.67	.08
11	12.0-12.4	12.25	63.0	11.80	.45
18	11.5-11.9	11.75	64.0	11.90	-.15
32	11.0-11.4	11.25	58.8	11.38	-.13
16	10.5-10.9	10.75	57.5	11.25	-.50
18	10.0-10.4	10.25	44.0	9.90	.35
10	9.5-9.9	9.75	50.0	10.50	-.75
8	9.0-9.4	9.25	36.3	9.13	.12
5	8.5-8.9	8.75	28.3	8.33	.42
3	8.0-8.4	8.25	27.5	8.25	.00
3	7.5-7.9	7.75	25.0	8.00	-.25

The original data* from which Table VI was obtained are represented graphically in Figure 3. The coordinates of each dot in this figure represent the mental ages of a pupil as determined by the two scales. The abscissa represents mental age as determined by the Illinois Examination. The ordinate represents the mental age as determined by the Binet Scale. The two instruments yield approximately the same mental ages for a

*Two additional cases are represented in the figure.

number of pupils. For others the difference is marked. It must be remembered that neither of the instruments is perfect and that when a difference exists between the two mental ages it is likely to be due to the fact that both rather than only the mental age yielded by the Illinois General Intelligence Scale involve errors.

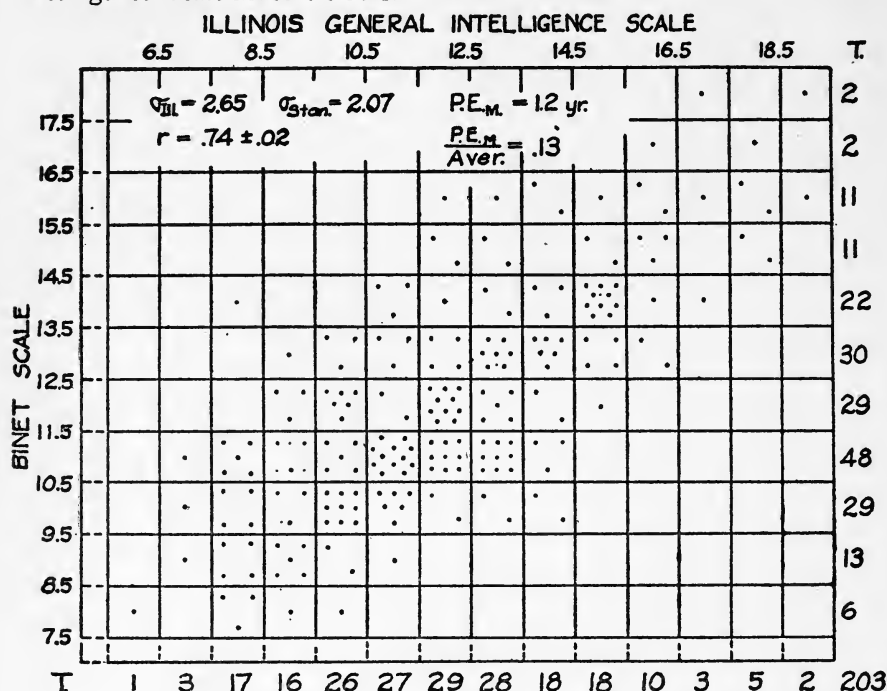


FIGURE 3. RELATION OF MENTAL AGES OF 203 PUPILS AS DETERMINED BY THE ILLINOIS GENERAL INTELLIGENCE SCALE AND BY THE BINET SCALE.

It should be borne in mind that mental ages above 14 obtained by the Illinois General Intelligence Scale must be interpreted somewhat differently from the mental ages obtained by other scales, especially the Stanford Revision of the Binet Scale. The extension of the concept of mental age which is here introduced is believed to be justified by the practical consideration of finding a convenient device to represent the absolute intelligence of bright children in the upper grades.

The intelligence quotient (I. Q.) is a measure derived by dividing a pupil's mental age by his chronological age. More strictly speaking the I. Q. is the quotient of a pupil's mental age divided by the median mental age for his chronological age. Below the age of 14 these are identical. A pupil's I. Q. is an index of his intelligence. If it is 100 he possesses only

average intelligence; if it is greater than 100, he is brighter than the average; if less than 100 he is duller. Because the mental ages above 14 do not have the same meaning as the corresponding mental ages obtained from the use of the Stanford Revision of the Binet Scale, it follows that the I. Q.'s for these ages also have a modified significance. For this reason Otis has called the quotient which we are using the "coefficient of brightness." Since below the age of 14 the quotient yielded by the Illinois General Intelligence Scale is identical in meaning with the quotient derived by the Stanford Revision of the Binet Scale, it has seemed wise to retain the name, intelligence quotient.

To facilitate the calculation of the I. Q. a table has been prepared which gives the I. Q.'s for each half year. This is reproduced as Table VII. The quotients have been calculated with reference to the mid-point of the intervals. Thus the quotient for chronological age group, 12-6, and mental age, 16-0, was found by dividing 16.25 by 12.75. A slightly different procedure was followed for chronological ages between 14 and 18. As noted above it is assumed that within this period there is a gradual slowing down of the growth of general intelligence. Hence, the differences between the actual median point scores for successive chronological age groups will gradually decrease and become zero at 18. Although the intelligence quotient is generally defined as the quotient obtained by dividing a pupil's mental age by his chronological age it is essentially the quotient obtained by dividing his mental age by the mental age norm for his chronological age. Below the age of 14 the norm and the pupil's chronological age are numerically the same. Above the age of 14 they are different. The quotients in Table VII for the age of 14 and above were calculated by using the mental age norm instead of the chronological age as the divisor.

The first column of the table contains point scores corresponding to the mental ages (M. A.'s) given in the second column. For example, a point score of 60 corresponds to a mental age of 11 years and six months, and a point score of 65 is equivalent to a mental age of 12 years. If a pupil's point score is 61, 62, 63, or 64, it is taken, when using this table, as meaning a mental age of 11 years and six months. Similarly, a pupil's chronological age is considered to be 10 years until it is 10 years and six months.

TABLE VII.-INTELLIGENCE QUOTIENTS*

Point Score	Mental Age	Chronological Age																	
		5-0	5-6	6-0	6-6	7-0	7-6	8-0	8-6	9-0	9-6	10-0	10-6	11-0	11-6	12-0	12-6	13-0	13-6
185	24-0														170	166	163	161	158
180	23-6														173	167	162	160	157
																	155	154	152
175	23-0														169	163	159	156	154
170	22-6														172	165	160	155	153
165	22-0														175	168	162	156	152
																	150	147	144
160	21-6														171	164	158	148	144
155	21-0														173	167	160	155	149
150	20-6														169	163	157	151	146
																	143	140	139
145	20-0														180	172	165	159	153
140	19-6														176	168	161	155	149
135	19-0														179	171	164	157	151
																	145	140	135
130	18-6														174	167	160	153	147
125	18-0														178	170	162	155	149
120	17-6														173	165	158	151	145
																	139	134	129
115	17-0														177	168	160	153	147
110	16-6														181	172	163	156	149
105	16-0														176	167	159	151	144
																	137	131	126
100	15-6														180	170	162	154	147
95	15-0														174	165	156	149	142
90	14-6														179	169	159	151	144
																	137	131	126
85	14-0														184	173	163	154	146
80	13-6														177	167	157	149	141
75	13-0														183	171	161	151	143
																	136	129	123
70	12-6														176	165	155	146	138
65	12-0														181	169	158	148	140
60	11-6														174	162	152	142	134
																	127	121	115
55	11-0														180	167	155	145	136
50	10-6														172	159	148	139	130
45	10-0														178	164	152	141	132
																	124	117	111
40	9-6														186	170	156	144	134
35	9-0														176	161	148	137	127
30	8-6														167	152	140	130	120
																	113	106	100
25	8-0														157	143	132	122	114
20	7-6														148	134	124	115	107
15	7-0														138	126	116	107	100
																	94	88	83
10	6-6														129	117	108	100	93
5	6-0														119	109	100	93	86
0	5-6														110	100	92	85	79
																	74	70	66
																	62	59	56
																	51	49	47
																	45	43	42
																	37	35	34
																	31	29	28
																	25	23	22
																	19	17	16
																	13	11	10
																	7	5	4
																	1	0	0

*Above the chronological age of 14, the Quotients are Coefficients of Brightness.

To find a pupil's I. Q. in this table proceed as follows: Take, for example, a pupil whose point score is 82 and whose chronological age is 11 years and four months. Find the interval* in the first column of numbers which contains a point score of 82. This is the 80 interval and the corresponding mental age (M. A.) is 13 years and 6 months. Locate at the top of the table the interval for a chronological age of 11 years; follow down the column for this interval to the line for the point score of 80. The number at the intersection of this column and this line is the pupil's I. Q. It is 122. *If a pupil's chronological age is more than 18 years, call it 18.* His I. Q. will always be found in the last column of the table.

As we shall show later there is evidence that the intelligence quotients yielded by this test have a greater variability than those obtained from the Stanford Revision of the Binet Scale. Therefore, when interpreting the I. Q.'s derived from the Illinois General Intelligence Scale in terms of degrees of brightness, it will be necessary to use a basis different from that proposed by Terman and others. (See page 67).

The distributions of the intelligence quotients furnish additional evidence of the validity of the rule for translating the point scores into mental ages. In Figure 4, the distribution of the I. Q.'s of the pupils of Decatur is represented graphically. In forming this distribution all grades have been combined, but the shape of the distributions for the separate half grades approximates that of this total distribution. The total number of pupils is 3787. The only conspicuous departure from the normal curve is in the

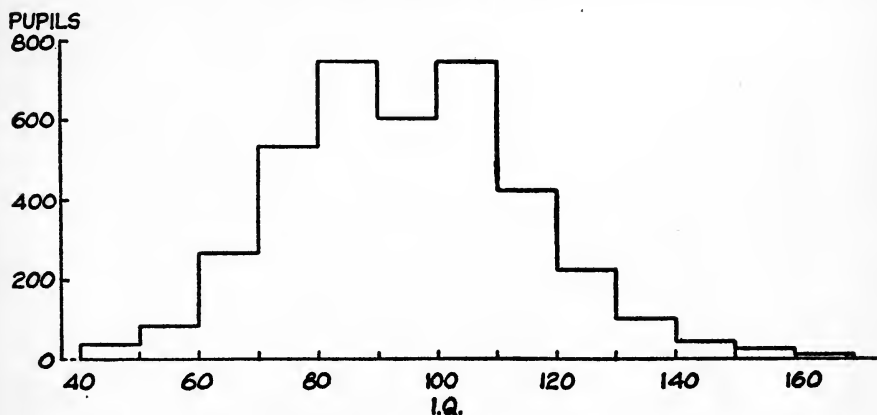


FIGURE 4. DISTRIBUTION OF PUPILS ACCORDING TO THEIR INTELLIGENCE QUOTIENTS.

*The term "interval" is used to indicate a span of point scores or ages. For the intelligence scores the span is five scores and in the table only the lower limit of this interval or span is given. For ages the interval is six months, and again only the lower limit is given.

90-100 interval. This is due to the structure of the table used for calculating I. Q.'s. Thus, we may say that the distributions of I. Q.'s do not indicate any error resulting from the rule used for translating point scores into mental ages.

Derived scores for achievement tests. To facilitate the comparison of a pupil's achievement with his mental age which represents his capacity to learn, provision was made for translating the point scores derived from the tests on arithmetic and silent reading into achievement age scores. The method of arriving at the basis for making this translation is similar to that used in the case of an intelligence scale. The point scores derived from all grades were thrown together and a distribution of the achievement scores was secured for each mental age group.* Because the data collected in May, 1920, were not representative, it is again necessary to illustrate the procedure and then to give the estimates which are recommended for use. Table VIII gives the median achievement point scores for the mental age groups for the city of Decatur. In this table the interval of mental

TABLE VIII. MEDIAN ACHIEVEMENT POINT SCORES FOR MENTAL AGE GROUPS IN DECATUR

Mental Age Interval	Average Mental Age	Number of Pupils	Arithmetic		Silent Reading Comprehension		Silent Reading Rate	
			Median Point Score	Corresponding Age Score	Median Point Score	Corresponding Age Score	Median Point Score	Corresponding Age Score
18.0-18.9	18.5	13	78.4	16.3	17.7	18.2	233.2	21.2
17.0-17.9	17.5	26	57.6	13.7	15.5	16.0	206.0	17.5
16.0-16.9	16.5	62	63.2	14.4	15.5	16.0	218.0	19.0
15.0-15.9	15.5	146	62.4	14.3	14.4	14.9	202.8	17.1
14.0-14.9	14.5	211	56.0	13.5	13.9	14.4	185.2	14.9
13.0-13.9	13.5	323	53.2	12.9	13.0	13.5	170.0	13.0
12.0-12.9	12.5	432	48.0	12.5	12.0	13.0	157.6	11.8
11.0-11.9	11.5	485	40.8	11.7	10.0	11.0	148.0	11.0
10.0-10.9	10.5	481	34.4	10.8	9.7	10.7	139.4	10.4
9.0-9.9	9.5	479	24.0	9.5	8.2	9.6	124.0	9.5
8.0-8.9	8.5	487	16.0	8.6	6.6	8.8	105.2	8.7
7.0-7.9	7.5	373	10.0	8.0	4.4	8.1	88.0	8.1
6.0-6.9	6.5	251	6.0	7.6	3.0	7.5	70.6	7.6
5.0-5.9	5.5	23	5.0	7.5	.5	6.9	36.0	6.9

*Before doing this the achievement point scores derived from the scales designed for the upper sequence of grades were corrected by adding the numbers given on page 19.

age is one year.* There are 432 pupils whose mental ages fall between 12 and 13 years. Their average mental age is 12.5. Their median point score on the arithmetic scale is 48.0. This corresponds to an achievement age of 12.5. That is, the median achievement of the pupils belonging to a mental age group is taken as the equivalent of the corresponding achievement age. In comprehension of silent reading the median point score of these pupils was 12.0. This is taken as corresponding to the achievement age of 13.0.** In rate of silent reading their median point score was 157.6. This is taken as corresponding to an achievement age of 11.8. The lack of agreement between the achievement age in rate of silent reading and the average mental age of this group is due to the fact that this group has not made as high a score in rate of reading as pupils of this mental age normally do. It must be remembered that these tables are based upon the scores of one city only. In addition to the data secured from Decatur, it was possible to make use of the grade norms derived for the Illinois Examination from 49,500 scores. Table IX gives the median point scores and the median mental age for each grade.

TABLE IX. GRADE NORMS FOR ILLINOIS EXAMINATION

Grade	Median Mental Age	Arithmetic	Silent Reading	
			Comprehension	Rate
III	7.9	10	3.8	82
IV	9.4	21	7.7	122
V	10.7	35	9.8	142
VI	12.0	44	11.0	158
VII	13.1	53	12.1	170
VIII	14.3	60	13.5	183

*A mental age interval of six months is preferable. It is not used in this illustration because the data were originally tabulated in yearly intervals for another purpose.

**The achievement age scores corresponding to the comprehension point scores were obtained by means of Table X. The unit of the point score is relatively large. Therefore, precise statements of equivalence were not shown in this table because fractional point scores are never obtained for individual pupils. This accounts for some of the apparent irregularities in the age scores for comprehension of silent reading.

In Figure 5, the median point scores derived from the arithmetic scale in Decatur are represented graphically by small crosses. The grade norms are represented by small circles. In arriving at the general relationship which exists between achievement and mental age it is necessary to bear in mind that the groups in the extremes of Table VIII are small and perhaps not representative. The pupils whose mental ages fall between 5 and 6 years probably have not learned to read well. Since the intelligence scale requires that the pupil be able to read, these mental ages are probably lower than they should be. At the upper end of the table we, of course, have very bright pupils, and bright pupils are likely to make higher scores than average or dull pupils of the same age. In Figure 5, a line has been drawn which is judged to represent the relation between achievement in the operations of arithmetic and mental age. This line can be used as a means for translating point scores into achievement ages. No simple rule can be stated as in the case of the intelligence scale. The achievement ages corresponding to the point scores yielded by the arithmetic scale are given in Table X.

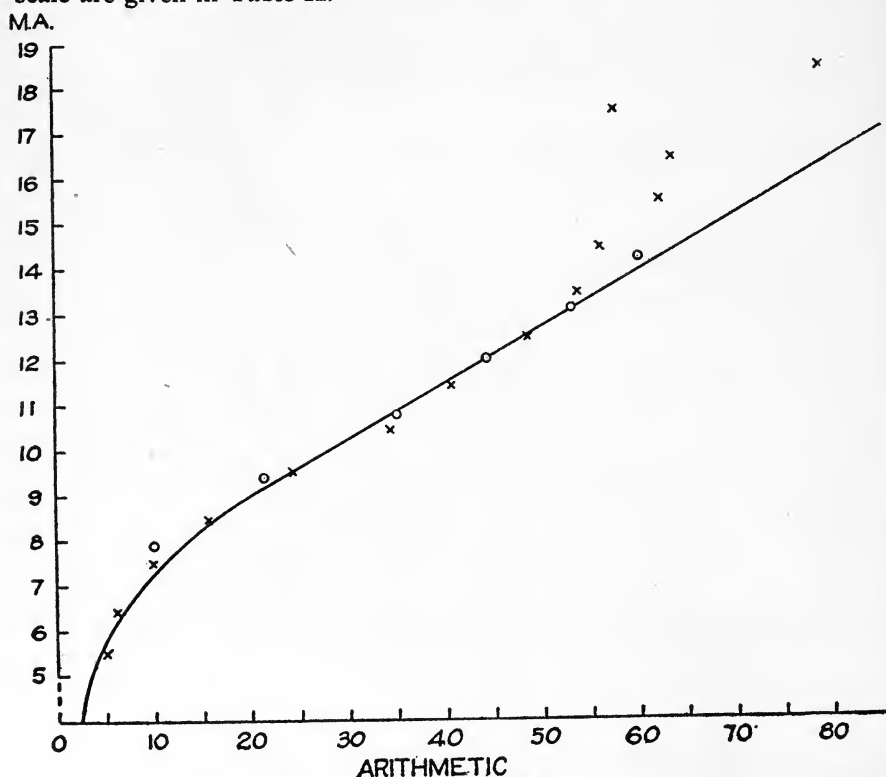


FIGURE 5. LINE OF RELATIONSHIP BETWEEN MEDIAN POINT SCORES IN ARITHMETIC AND MENTAL AGE.

TABLE X. CORRESPONDING ACHIEVEMENT AGE FOR POINT SCORES IN ARITHMETIC AND SILENT READING

Rate	Point scores		Achievement Age
	Comprehension	Arithmetic	
250		140	24-0
247		136	23-6
244		132	23-0
241		128	22-6
238		124	22-0
235		120	21-6
232	20	116	21-0
229		112	20-6
226		108	20-0
222	19	104	19-6
218		100	19-0
214	18	96	18-6
210		92	18-0
206	17	88	17-6
202		84	17-0
198	16	80	16-6
194		76	16-0
190	15	72	15-6
186		68	15-0
182	14	64	14-6
178		60	14-0
174	13	56	13-6
170	12	52	13-0
165		48	12-6
160	11	44	12-0
154		40	11-6
148	10	36	11-0
141		32	10-6
133	9	28	10-0
124	8	24	9-6
113	7	20	9-0
100	6	15	8-6
85	4	10	8-0
67	3	5	7-6
47	1	0	7-0
25	0		6-6
0			6-0

Figure 6 and Figure 7 represent corresponding data for comprehension and rate of silent reading. The point scores corresponding to the various achievement ages are also given in Table X.

When measures of both achievement and general intelligence are expressed in terms of ages they may be considered comparable. The Illinois Examination provides for comparison by dividing the pupil's achievement age by his mental age.* The quotient is called his achievement quotient.(A. Q.)

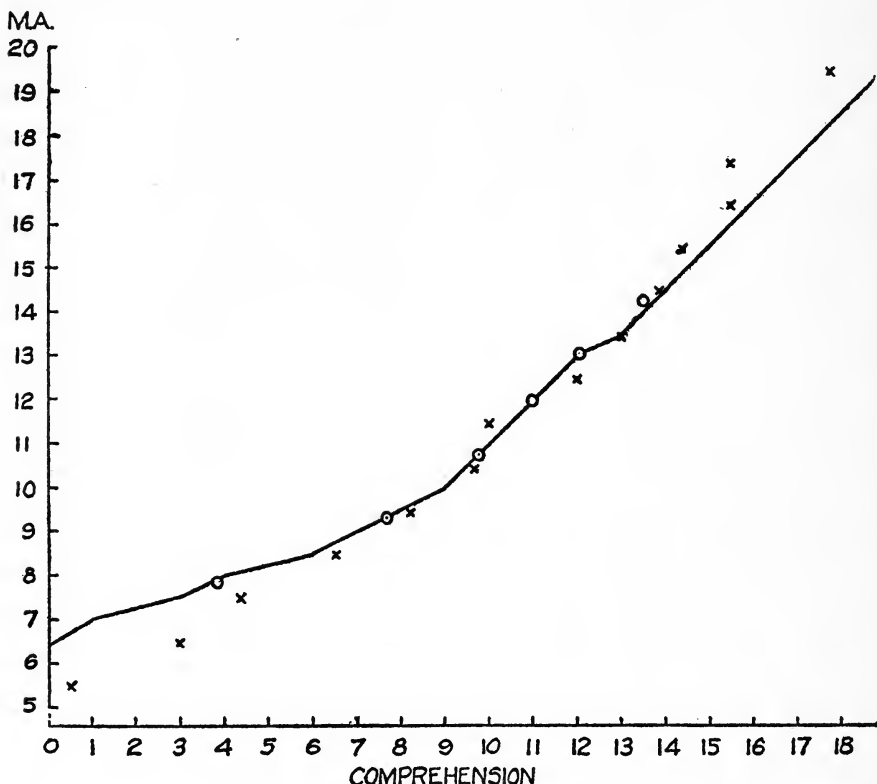


FIGURE 6. LINE OF RELATIONSHIP BETWEEN MEDIAN POINT SCORES IN COMPREHENSION OF SILENT READING AND MENTAL AGE.

*The achievement age norm for a pupil is numerically the same as his mental age. Hence, this procedure may also be thought of as being a comparison of a pupil's achievement with the norm for his mental age.

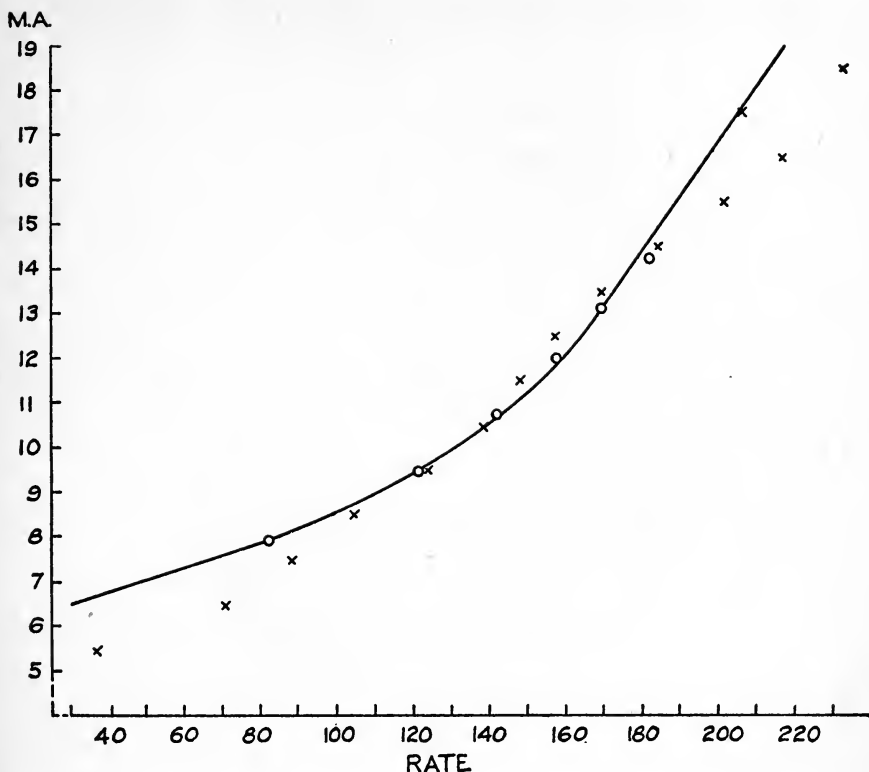


FIGURE 7. LINE OF RELATIONSHIP BETWEEN MEDIAN POINT SCORES IN RATE OF SILENT READING AND MENTAL AGE.

This is an index of the extent to which his achievement corresponds to his general intelligence or capacity to achieve. For example, if a pupil has an achievement age of 12 years and a mental age of 9 years, his achievement quotient is 133. (The decimal point is omitted as in the case of intelligence quotients.) If this pupil's mental age had been 14 years, his achievement quotient would have been 86. An achievement quotient of 100 means that the pupil has achieved exactly the average of pupils of his mental age, or that he is just up to the norm for his mental age. If his achievement quotient is 130, he has achieved thirty percent more than the average of the pupils of his mental age; on the other hand, if his achievement quotient is 75, we have evidence that he has achieved only seventy-five percent as much as the average of pupils of his mental age.

TABLE XI--ACHIEVEMENT QUOTIENTS

Point Scores			Achievement Age	Mental Age																			
Rate	Comprehension	Arithmetic		4-6	5-0	5-6	6-0	6-6	7-0	7-6	8-0	8-6	9-0	9-6	10-0	10-6	11-0	11-6	12-0	12-6	13-0	13-6	14-0
259		152	25-6										264	251	240	229	219	210	202	194	187	181	
256		148	25-0											259	246	235	224	215	206	198	191	183	177
253		144	24-6										254	241	230	220	211	202	194	187	180	174	
250	23	140	24-0										249	237	226	216	206	198	190	183	176	170	
247		136	23-6										244	232	221	211	202	194	186	179	173	167	
244	22	132	23-0										238	227	216	207	198	190	182	175	169	163	
241		128	22-6										260	246	233	222	212	202	194	186	178	172	
238		124	22-0										254	241	228	217	207	198	189	182	175	168	
235	21	120	21-6										249	235	223	212	202	193	185	178	171	164	
232		116	21-0										243	230	218	207	198	189	181	173	167	160	
229		112	20-6										231	217	204	193	184	177	169	163	157	151	
226	20	108	20-0										245	231	219	208	198	188	180	172	165	159	
222		104	19-6										239	226	214	203	193	184	176	168	161	155	
218	19	100	19-0										233	220	208	197	188	179	171	164	157	151	
214	18	96	18-6										227	214	203	192	183	174	167	160	153	147	
210		92	18-0										221	209	197	187	178	170	162	155	149	143	
206	17	88	17-6										205	192	182	173	165	158	151	145	139	134	
202		84	17-0										197	186	177	168	160	153	147	141	135	130	
198	16	80	16-6										191	181	172	163	156	149	143	137	131	126	
194		76	16-0										180	170	162	154	147	140	134	129	124	119	
190	15	72	15-6										173	163	154	146	139	133	127	121	116	112	
186		68	15-0										165	156	149	142	136	130	124	120	115	111	
182	14	64	14-6										159	151	144	137	131	126	120	116	111	107	
178		60	14-0										154	146	139	133	127	121	116	112	108	104	
174	13	56	13-6										149	141	134	128	122	117	112	108	104	100	
170	12	52	13-0										136	129	123	118	113	108	104	100	96	93	
165		48	12-6										243	221	204	189	176	165	155	146	138	131	
160			12-0										263	239	220	204	189	177	167	157	149	141	
154		40	11-6										252	230	212	196	183	171	161	151	143	136	
148	10	36	11-0										247	228	211	196	184	173	163	154	146	139	
141			10-6										233	213	196	181	169	158	148	140	132	126	
133			10-0										247	204	204	188	174	162	152	142	134	127	
124			9-6										237	214	195	180	167	155	145	136	129	122	
113	9	32	9-0										226	205	187	172	159	148	139	130	123	116	
100	8	28	8-6										216	195	178	164	152	141	132	124	117	111	
85	4	15	8-0										205	186	170	156	144	134	126	118	111	105	
	7	20	7-6										195	176	161	148	137	127	119	112	106	100	
	6	15	7-0										184	167	152	140	130	120	113	106	100	95	
	0	10	6-6										174	157	143	132	122	114	106	100	94	89	
67	5	5	7-6										163	148	134	124	115	107	100	94	89	84	
47	3	0	7-0										153	138	126	116	107	100	94	88	83	78	
25	0		6-6										142	129	117	108	100	93	87	82	77	73	
			6-0										131	119	109	100	93	86	81	76	71	68	
													4-6	5-0	5-6	6-0	6-6	7-0	7-6	8-0	8-6	9-0	
																				</			

Rate	Point Scores		Achievement Age	Mental Age																23-6	24-0
	Comprehension	Arithmetic		14-6	15-0	15-6	16-0	16-6	17-0	17-6	18-0	18-6	19-0	19-6	20-0	20-6	21-0	21-6	22-0		
259		152	25-6	175	169	163	158	154	149	144	141	137	134	130	127	124	121	118	116	113	108
256		148	25-0	171	166	160	155	151	146	142	139	135	131	127	125	122	119	116	114	111	106
253		144	24-6	167	162	157	152	148	143	139	136	132	129	125	122	119	116	114	111	109	104
250	23	140	24-0	164	159	154	149	145	141	137	133	129	126	123	120	117	114	112	109	107	102
247		136	23-6	161	156	151	146	142	138	134	130	127	123	120	117	114	112	109	107	104	98
244	22	132	23-0	157	152	148	143	139	135	131	127	124	121	118	115	112	109	107	105	102	96
241		128	22-6	154	149	144	140	136	132	128	125	121	118	115	112	110	107	105	102	100	94
238	21	124	22-0	151	146	141	137	133	129	125	122	119	116	113	110	107	105	102	100	98	92
235		120	21-6	147	143	138	134	130	126	123	119	116	113	110	107	105	102	100	98	96	90
232		116	21-0	144	139	135	131	127	123	120	116	113	110	108	105	102	100	98	96	93	88
229	20	112	20-6	141	136	132	128	124	120	117	114	111	108	105	102	100	98	95	93	91	85
226		108	20-0	137	133	129	125	121	117	114	111	108	105	103	100	98	95	93	91	89	84
222	19	104	19-6	134	130	125	122	118	114	111	108	105	103	100	98	95	93	91	89	87	81
218	18	100	19-0	130	126	122	118	115	112	108	105	103	100	97	95	93	90	88	86	84	79
214		96	18-6	127	123	119	115	112	109	106	103	100	97	95	92	90	88	86	84	82	77
210		92	18-0	124	120	116	112	109	106	103	100	97	95	92	90	88	86	84	82	80	75
206	17	88	17-6	120	116	113	109	106	103	100	97	95	92	90	88	86	84	82	80	78	73
202		84	17-0	117	113	110	106	103	100	97	94	92	90	87	85	83	81	79	78	76	71
198	16	80	16-6	114	110	106	103	100	97	94	92	89	87	84	82	80	78	76	75	73	69
194		76	16-0	110	107	103	100	97	94	91	88	86	84	82	80	78	76	74	72	71	67
190	15	72	15-6	107	103	100	97	94	91	88	86	84	82	80	78	76	74	72	71	69	65
186		68	15-0	103	100	97	94	91	88	86	84	81	79	77	75	73	72	70	69	67	64
182	14	64	14-6	100	97	94	91	88	86	83	81	79	77	75	73	71	69	68	66	65	63
178		60	14-0	97	93	90	88	85	83	80	78	76	74	72	70	69	67	66	64	63	59
174	13	56	13-6	93	90	87	85	82	79	77	75	73	71	69	67	65	63	62	60	59	57
170	12	52	13-0	90	87	84	82	79	77	75	73	71	69	67	65	63	62	61	60	58	55
165		48	12-6	86	84	81	78	76	74	72	70	68	66	65	63	61	60	58	57	56	52
160	11	44	12-0	83	80	78	75	73	71	69	67	65	64	62	61	59	58	56	55	54	50
154		40	11-6	80	77	75	72	70	68	66	64	63	61	60	58	57	55	54	53	51	48
148	10	36	11-0	76	74	71	69	67	65	63	62	60	58	57	55	53	52	51	49	48	46
141		32	10-6	73	70	68	66	64	62	61	59	57	56	54	53	52	51	49	48	47	45
133	9	28	10-0	69	67	65	63	61	59	58	56	55	53	52	51	49	48	47	46	45	44
124	8	24	9-6	66	64	62	60	58	57	55	53	52	51	49	48	47	46	45	44	43	41
113		20	9-0	63	61	59	57	55	54	52	51	49	48	47	46	45	44	43	42	41	39
100	6	15	8-6	59	57	56	54	52	51	49	48	47	45	44	43	42	41	40	39	38	36
85	4	10	8-0	56	54	52	51	49	48	46	45	44	43	42	41	40	39	38	37	36	34
67	3	5	7-6	53	45	49	48	46	45	44	42	41	40	39	38	37	36	35	34		
47	1	0	7-0	49	48	46	45	44	43	42	41	40	39	38	37	36	35	34			
25	0		6-6	46	44	43	42	40	39	38	37	36	35	34							
0			6-0	42	41	40	38	37	36	35	35	34	33	32	31	30	29	28	27	26	24-0

Mental Age

Table XI is to be used in obtaining the achievement quotients. Take, for example, a pupil who has a rate score of 165, a comprehension score of 10, and an intelligence score of 60. An intelligence score of 60 is equivalent to a mental age of 11-6. (See Table VII). A rate score of 165 is equivalent to an achievement age of 12-6, and a comprehension score of 10 is equivalent to an achievement age of 11-0. To find the achievement quotient for rate of silent reading, find at the left of the table the line for an achievement age of 12-6 years. At the top of the table, find the column for a mental age of 11-6; follow down this column until you come to the line for 12-6. The number at this intersection point is the achievement quotient for rate. It is 109. To find the achievement quotient for comprehension, follow down the same column until you come to the line for an achievement age of 11-0; the number at this intersection, 96, is the achievement quotient for comprehension.

Mental age norms versus chronological age norms. In determining a basis for translating point scores into age scores we have grouped the pupils according to their mental ages as given by the Illinois General Intelligence Scale. A grouping according to chronological age would have been more convenient. Theoretically, the same results should have been obtained because the average mental age of an unselected chronological age group is identical with the average chronological age of that group. The determining motive for devising a procedure for translating point scores into age scores was to secure a convenient method for obtaining an achievement quotient. This quotient expresses the relation between a pupil's achievement and his individual achievement norm. A pupil's capacity to learn or his mental age is a much more potent factor in determining his achievement than is his chronological age. In interpreting achievement we must compare it with a pupil's capacity to learn. Hence, logically, a grouping of pupils according to mental age is to be preferred. However, since the same numerical results will be obtained from a chronological age grouping there is practical justification for using it because of the greater convenience.

Justification of disregarding school grade in determining achievement ages. In determining the median point score of pupils having the same mental age no attention was given to the school grade to which the pupils belong. Pupils in the third grade were grouped with pupils in the eighth, providing they had the same mental age. This procedure implies that in judging a pupil's achievement his school grade may be disregarded. In other words, his achievement depends only or largely upon his mental age and upon the effectiveness of the instruction received, and not on the quantity of the instruction as measured by the grade attained. The validity of this assumption was investigated by calculating separately for each

school grade the median achievement age scores for each mental age group. Figure 8 represents graphically these median achievement age scores for arithmetic in the Decatur schools. It will be noticed here that there is, in general, a distinct increase from grade to grade in the median point score of pupils belonging to the same mental age. Thus, in this case our assumption is not in entire agreement with the facts. However, convenience of use demands that we have only one rule for translating point scores into achievement ages. If a different one were used for each half grade or for each grade the inconvenience and confusion arising would decidedly limit the usefulness of the scale. Furthermore, any injustice which may be done to pupils in the lower grades may be compensated for in the interpre-

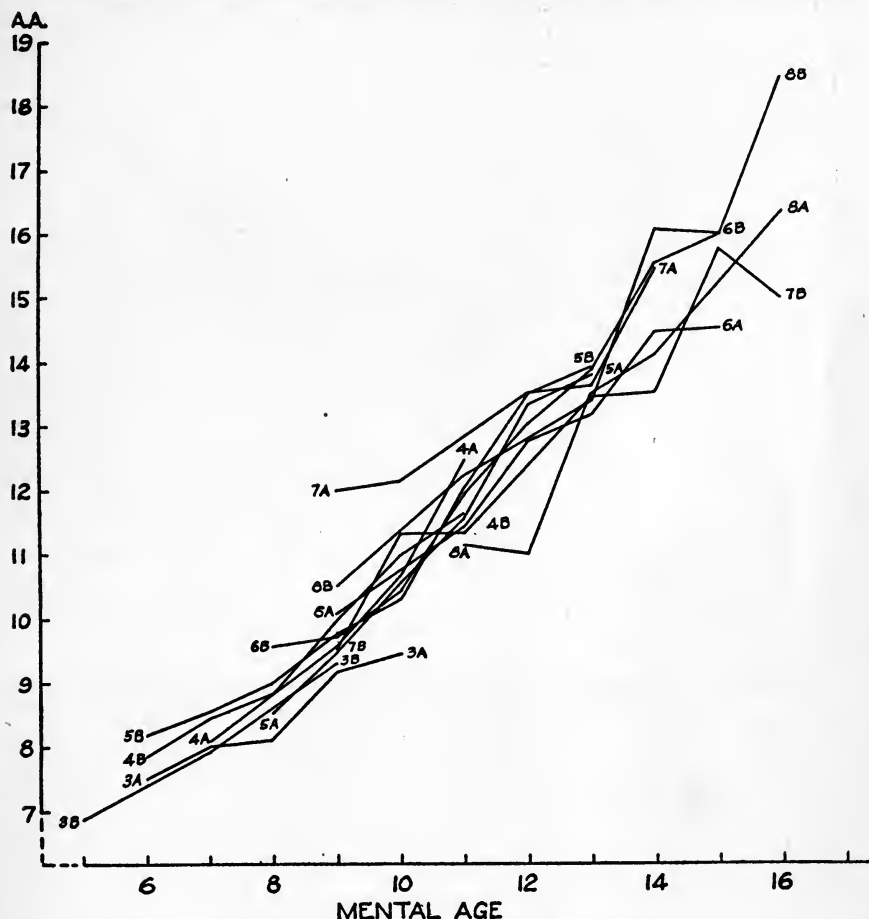


FIGURE 8. EFFECT OF SCHOOL GRADE UPON ACHIEVEMENT IN ARITHMETIC.

tation of the achievement quotients. The grade norms, however, show that in general the increase from grade to grade is slight. (See page 66).

Figures 9 and 10 report similar data for comprehension and rate of silent reading. In both of these figures it is practically impossible to distinguish between the lines representing the median achievements in the lower grades and those representing the median achievements in the upper grades. Consequently, we may say that our assumption in the case of silent reading that we disregard the placement of the pupil in translating his point score into an achievement age score is closely in agreement with the facts. Another interpretation of these figures is that ability to read silently correlates highly with general intelligence.

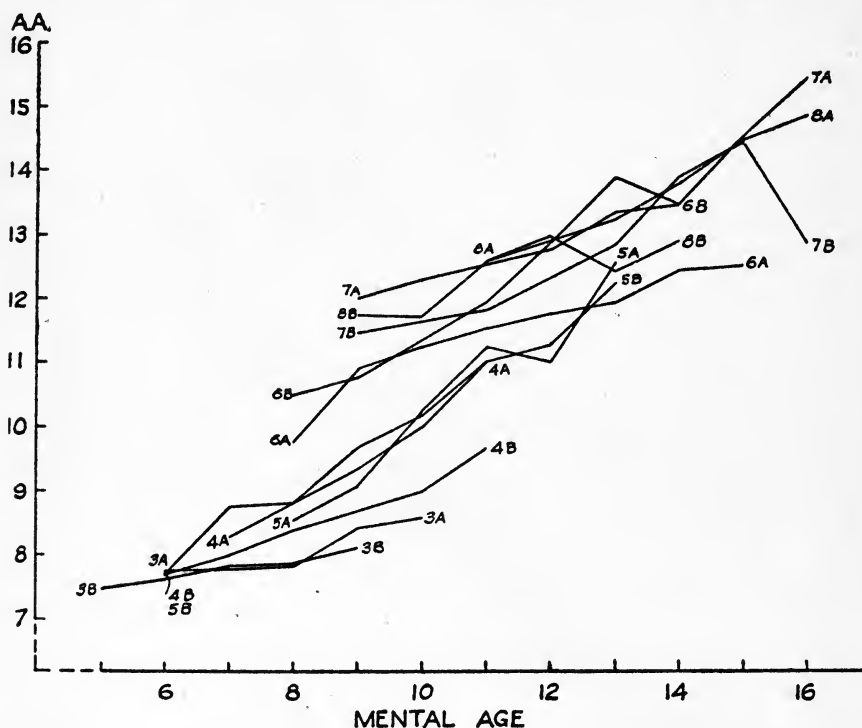


FIGURE 9. EFFECT OF SCHOOL GRADE UPON ACHIEVEMENT IN COMPREHENSION OF SILENT READING.

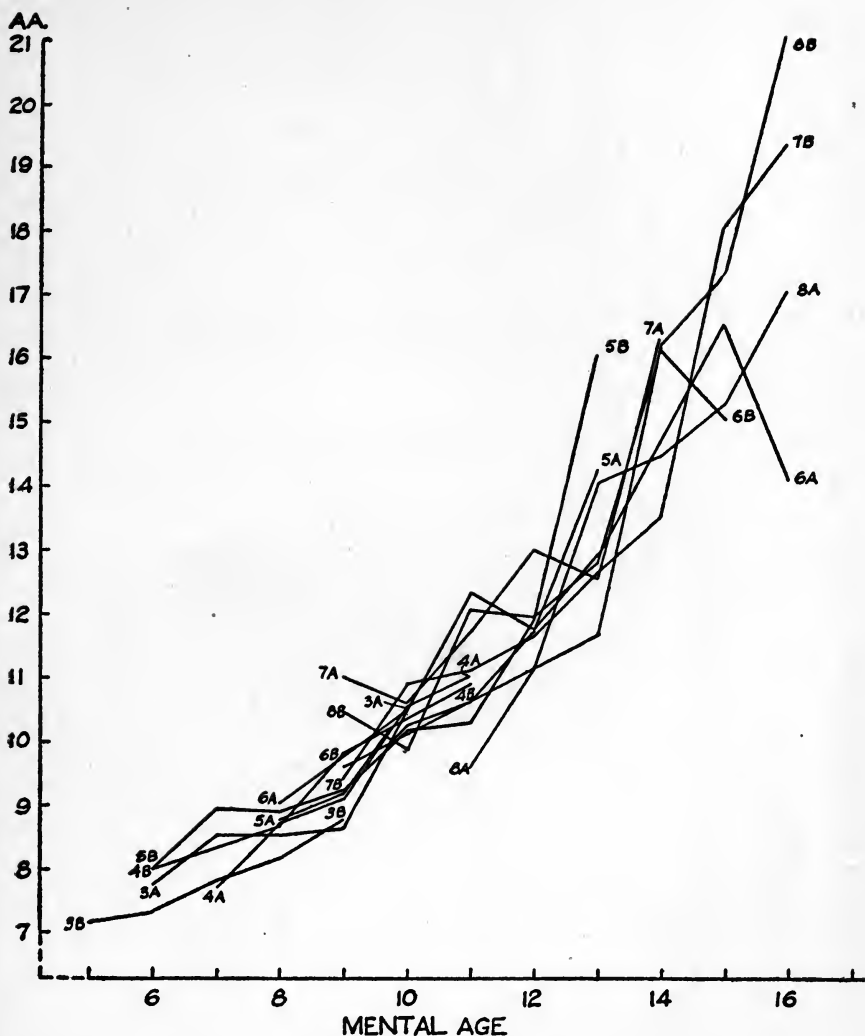


FIGURE 10. EFFECT OF SCHOOL GRADE UPON ACHIEVEMENT IN RATE OF SILENT READING.

The Illinois Examination yields four age scores and four quotient scores for each pupil. Provision has been made to calculate an average score in the case of silent reading. When this is done each pupil has ten scores. Table XII gives representative scores.

TABLE XII. TYPICAL SCORES

Pupil	Grade	Ages				Quotients			
		Mental	Arith- metic	Silent Reading		Intelli- gence	Arith- metic	Silent Reading	
				Compre- hension	Rate			Compre- hension	Rate
1	8A	12-0	10-0	11-0	11-0	83	84	92	92
2	8B	12-0	12-0	9-6	9-0	89	100	80	76
3	5A	9-0	8-0	9-0	8-0	90	89	100	89
4	3B	7-6	7-6	8-6	7-6	89	100	113	100
5	3A	6-6	7-0	6-6	6-0	73	107	100	93
6	6A	12-0	11-0	9-6	14-0	100	92	80	116
7	6B	11-6	12-0	13-0	11-6	100	104	113	100
8	4A	8-0	8-0	10-0	10-6	73	100	124	130
9	4B	7-6	7-6	8-0	7-6	79	100	106	100
10	5B	11-0	10-6	13-0	10-6	115	96	118	96
11	7A	12-6	14-0	9-6	10-6	93	112	76	84
12	7B	15-0	15-0	22-0	21-0	124	100	146	139

IV. FUNCTION.

The function of the three scales which make up the Illinois Examination is implied in their structure. The Illinois General Intelligence Scale provides a measure of general intelligence of children in grades III to VIII, inclusive.* Monroe's Standardized Silent Reading Tests, Revised, are intended to yield measures of the ability to read silently simple descriptive and narrative material when the reading is done for the purpose of answering questions. Monroe's General Survey Scale in Arithmetic is designed to yield general measures of a pupil's ability to perform the operations of arithmetic. It should be noted that the function of these tests is general rather than diagnostic. It is possible to use the sub-tests of the General Survey Scale in Arithmetic as diagnostic tests although they were not designed for this purpose.

The function of the Illinois Examination as a whole is indicated in the introductory statement.

V. VALIDITY

The ideal procedure to be followed in studying the truthfulness of the measures yielded by the scales included in the Illinois Examination would be to compare them with true measures secured by other means. However, in no case are such true measures available. It is, therefore, necessary to study the validity of these scales by methods which are obviously imperfect.

1. **Objectivity.** The scales of the Illinois Examination are highly objective with respect to the scoring of test papers. Except when a pupil fails to follow directions no questions, concerning which answers are correct, arise. The administration of the tests is also highly objective. The directions for examiners have been found to be adequate and in all cases the examiner is told very explicitly what he is to say to the pupils. Much of the explanation is also printed on the test booklet so that the pupil has an opportunity to read as well as to hear the explanation.

2. **Reliability**.** In order to study the reliability of the three scales which make up the Illinois Examination the different forms were given to the same pupils. The instruction to those cooperating in this study was to give all the forms within the same half day. The scores on the different forms were compared by means of the Pearsonian coefficient of correlation and by means of other statistical devices which will be explained in the following pages.

*It is also recommended for use in the high school.

**A related phase of reliability is considered on page 68 under the head of "Practice effect when test is repeated."

The coefficient of correlation merely indicates the relationship between two sets of scores. It is simply an index of the extent to which the pupils make the same score on the second trial of a test that they make upon the first trial when the practise effect is disregarded. In Figure 11, we represent graphically the scores made by the fifth grade pupils on Forms 1 and 2 of the Illinois General Intelligence Scale. The coefficient of correlation for this group of scores is $.92 \pm .006$. It is obvious that in some cases the pupils make the same score or approximately the same score on second trial. In other cases there are marked differences between the scores on the two trials.

In Figure 11 the regression line, $y = 4.92 + .80 x$, has been drawn. Perfect correlation ($r_{12} = 1.00$) would be secured if the Form 1 scores were changed so that all points would fall upon this regression line. This would require a vertical shifting of the points. Those above would be moved downward, while those below would be moved upward. For a few of the points vertical lines have been drawn in to indicate the amount of shifting necessary. Perfect correlation would be secured if this were done with reference to any line but this regression line is the one for which the standard deviation of the shifting is the least. The other regression equation, $x = 4.69 + 1.05 y$, possesses similar properties for a horizontal shifting.

The amounts of changes necessary to secure perfect correlation may be thought of as departures from perfect correlation. The magnitude of these changes is described by the equation for the probable error of estimate,

$$P. E._{est} = .6745 \sigma_y \sqrt{1 - r_{12}^2}.$$

Substituting in this equation for σ_y and r_{12} we have the probable error of estimate equal to 6.06. The probable error of estimate is more easily interpreted as the index of the degree of correlation that exists than as the coefficient of correlation.

Since neither set of scores gives accurate measures of intelligence the differences between the pairs of scores do not truthfully represent the degree of inaccuracy of either set of scores. The error of any score is the difference between it and a pupil's true score. We may define a true score as the average of an infinite number of scores after they have been corrected for practise effect, fatigue, and other factors which would tend to increase or decrease the averages of the successive sets of scores. Such true scores are obviously not obtainable. It is, however, possible to determine the coefficient of correlation between either set of obtained scores and the

corresponding true scores. This is done by the formula,*

$$r_{1t} = \sqrt{r_{12}}$$

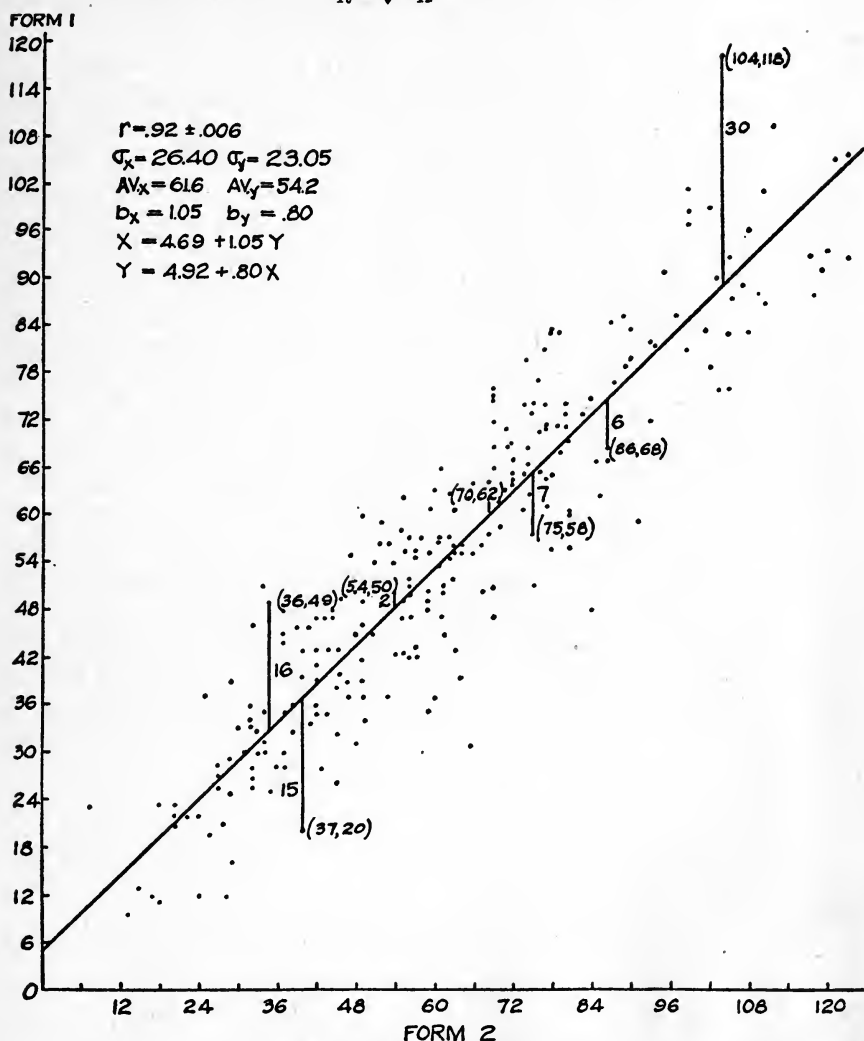


FIGURE 11. CORRELATION OF FORM 1 SCORES WITH FORM 2 SCORES OF THE ILLINOIS GENERAL INTELLIGENCE SCALE, FIFTH GRADE.

In this formula r_{12} is the coefficient of correlation between the two sets of obtained scores and r_{1t} is the coefficient of correlation between one set of obtained scores and the corresponding true scores. To distinguish the

*See Kelley, T. L. "A simplified method of using scaled data for purposes of testing," *School and Society*, 4:74, July 8, 1916, and "The reliability of test-scores," *Journal of Educational Research*, 3:370-79, May, 1921.

coefficient of correlation of a set of obtained scores with the corresponding set of true scores from the coefficient of correlation between two sets of obtained scores we call the latter the *coefficient* of reliability and the former the *index* of reliability.

Table XIII gives coefficients of reliability (r_{12}) for each of the three scales which make up the Illinois Examination. Table XIV gives the corresponding indices of reliability (r_{1t}). It will be noted that a high degree of reliability is indicated in most cases. In some instances it is unusually high in comparison with the degree of reliability reported for other tests. In the case of silent reading certain instructions were not followed by some of the examiners, and it is thought that their failure to do so caused the two sets of scores to correlate less highly than they should.

Both the coefficient of reliability and the index of reliability are difficult to interpret. They express a general relationship but not in terms of the actual amount of error which must be allowed for in the case of the scores of individual pupils. It is possible to calculate another and more easily interpreted expression of the reliability or accuracy of the measures yielded by tests. The probable error of estimate is given by the formula,

$$P. E._{est} = .6745 \sigma \sqrt{1 - r_{1t}^2}.$$

In the formula σ may be taken as either σ_1 or σ_2 . Theoretically, these are expected to be equivalent. Practically, slight differences may exist. It is, therefore, advisable to use as the value of σ the average of σ_1 and σ_2 . The probable error of estimate ($P. E._{est}$) in this case is essentially the probable error in the measurements yielded by the test. Hence, we call it by this name and use the symbol, $P. E._m$. Since $r_{1t} = \sqrt{r_{12}}$, the above formula may be written in the form,

$$P. E._m = .6745 \sigma \sqrt{1 - r_{12}}.$$

The probable error of measurement calculated by the above formula is to be interpreted as an index of the amount of departure of the obtained scores from the true scores. In other words, it is the error which the obtained score involves. This error is described as a probable error. Such a description, of course, tells us nothing about the magnitude of the error in the case of a particular pupil but it does describe in a general way the magnitude of the errors involved in a group of scores.

TABLE XIII. RELIABILITY COEFFICIENTS

Grade	Form	General Intelligence		Arithmetic		Silent Reading		
		No. of Pupils	r_{12}	No. of Pupils	r_{12}	No. of Pupils	Rate r_{12}	Comprehension r_{12}
III	1 with 2	76	.86	299	.80	116	.69	.63
	2 with 3						.72	.64
	1 with 3						.72	.63
IV	1 with 2	120	.93	271	.86	112	.79	.63
	2 with 3						.80	.64
	1 with 3						.70	.60
V	1 with 2	243	.92	256	.88	120	.79	.69
	2 with 3						.86	.74
	1 with 3						.72	.71
III to V	1 with 2			826	.95	348	.78	.80
	2 with 3						.85	.64
	1 with 3						.76	.82
VI	1 with 2	198	.82	271	.76	139	.87	.52
	2 with 3						.89	.86
	1 with 3						.90	.70
VII	1 with 2	157	.80	257	.71	100	.72	.68
	2 with 3						.74	.75
	1 with 3						.61	.63
VIII	1 with 2	164	.67	171	.79	119	.91	.85
	2 with 3						.88	.71
	1 with 3						.91	.85
VI to VIII	1 with 2			699	.76	358	.79	.72
	2 with 3						.84	.76
	1 with 3						.75	.76
III to VIII	1 with 2	958	.92					

TABLE XIV. INDEX OF RELIABILITY

Grade	Form	General Intelligence		Arithmetic		Silent Reading		
		No. of Pupils	r_{1t}	No. of Pupils	r_{1t}	No. of Pupils	Rate r_{1t}	Comprehension r_{1t}
III	1 with 2	76	.93	299	.89	116	.83	.79
	2 with 3						.85	.80
	1 with 3						.85	.79
IV	1 with 2	120	.96	271	.93	112	.89	.79
	2 with 3						.89	.80
	1 with 3						.84	.77
V	1 with 2	243	.96	256	.94	120	.89	.83
	2 with 3						.93	.86
	1 with 3						.85	.84
III to V	1 with 2			826	.97	348	.88	.89
	2 with 3						.92	.80
	1 with 3						.87	.91
VI	1 with 2	198	.91	271	.87	139	.93	.72
	2 with 3						.94	.93
	1 with 3						.95	.84
VII	1 with 2	157	.89	257	.84	100	.85	.82
	2 with 3						.86	.87
	1 with 3						.78	.79
VIII	1 with 2	164	.82	171	.89	119	.95	.92
	2 with 3						.93	.84
	1 with 3						.95	.92
VI to VIII	1 with 2			699	.87	358	.89	.85
	2 with 3						.92	.87
	1 with 3						.87	.87
III to VIII	1 with 2	958	.96					

TABLE XV. PROBABLE ERRORS OF MEASUREMENT AND RATIO OF PROBABLE ERRORS OF MEASUREMENT TO AVERAGE SCORES

Grade	Forms	Intelligence		Arithmetic		Silent Reading			
		P.E. _m	P.E. _m	P.E. _m	P.E. _m	Comprehension		Rate	
			Av.		Av.	P.E. _m	P.E. _m	P.E. _m	P.E. _m
							Av.		Av.
III	1 with 2	3.5	0.10	2.6	0.17	1.2	0.16	13.7	0.12
	1 with 3					1.3	0.17	13.7	0.11
	2 with 3					1.3	0.17	14.5	0.12
IV	1 with 2	5.5	0.09	4.6	0.12	1.4	0.16	10.3	0.08
	1 with 3					1.2	0.14	12.8	0.10
	2 with 3					1.2	0.13	10.9	0.09
V	1 with 2	4.7	0.08	4.4	0.09	1.2	0.10	12.0	0.08
	1 with 3					0.9	0.07	13.7	0.08
	2 with 3					1.1	0.09	10.7	0.07
III to V	1 with 2			3.2	0.10	1.0	0.11	13.1	0.10
	1 with 3					1.0	0.10	13.7	0.10
	2 with 3					0.9	0.09	11.5	0.08
VI	1 with 2	5.5	0.07	6.3	0.10	1.3	0.10	9.1	0.05
	1 with 3					1.1	0.09	8.2	0.05
	2 with 3					1.1	1.08	9.5	0.05
VII	1 with 2	6.4	0.07	5.3	0.09	1.2	0.09	13.6	0.07
	1 with 3					1.4	0.10	17.0	0.09
	2 with 3					1.1	0.07	14.9	0.07
VIII	1 with 2	7.7	0.08	5.4	0.08	0.8	0.05	7.5	0.04
	1 with 3					0.9	0.06	7.5	0.04
	2 with 3					1.1	0.07	9.8	0.05
VI to VIII	1 with 2			6.2	0.10	1.1	0.08	12.0	0.07
	1 with 3					1.1	0.08	13.7	0.07
	2 with 3					1.1	0.08	11.4	0.06
III to VIII	1 with 2	5.3	0.07						

The probable error of measurement tends to increase as the scores become larger and the significance of an error depends upon the magnitude of the score with which it is associated. For this reason added meaning can be given to our description of the errors by calculating the ratio of the probable error of measurement to the average score. This gives the probable error of measurement in the form of a percent of the score. In Table XV, the probable errors of measurement and the ratios of these to the average scores are given. The numbers of pupils involved are the same as those given in Tables XIII and XIV. A probable error of measurement of 3.5 for the Illinois General Intelligence Scale in the third grade means that the point scores of fifty percent of the pupils will involve errors less than this amount. The remaining fifty percent of the scores will involve errors greater than 3.5. A somewhat more general statement is that, on the average, the scores obtained from third grade pupils will involve probable errors of ten percent of their magnitude. The larger the score the larger the probable error.

The magnitude of these probable errors of measurement may appear to be somewhat disturbing but they are not large in comparison with those calculated for other tests. In the case of the Illinois General Intelligence Scale the average probable error of measurement amounts to about six months. This is approximately the same as that calculated for the Stanford Revision of the Binet Scale for the Measurement of Intelligence.* An unpublished study of the reliability of a group of silent reading tests made by the writer gave ratios of reliability which are much larger in many cases.

3. **Discrimination.** The shape of the distribution of the scores which a test yields throws some light upon its validity. In order that a test be valid the scores must show differences of the traits measured when these differences exist. When a representative group of pupils is measured with reference to a mental or a physical trait we may expect to find a distribution closely approximating the normal shape. When the number of cases is large this approximation should be close if there is proper discrimination. Any marked departure from the normal shape indicates that for some pupils at least there is a lack of discrimination. On the other hand, when we have a normal distribution we cannot know definitely that our measures are accurate. The shape of the distribution, therefore, has only a negative significance. We can only say that when there is a striking departure from normality there is a lack of discrimination and, hence, inaccuracy of measurement for some pupils.

*Otis, A. S. and Knollin, H. E. "Reliability of Binet Scale and Pedagogical Scales." *Journal of Educational Research*, 4:121-43, September, 1921.

Figures 12, 13, 14, and 15 represent graphically the distributions of scores yielded by the scales of the Illinois Examination for certain groups of pupils. In Figure 12, the distributions with respect to mental age are given for grades IV, VI, and VIII. Each of these distributions includes several thousand children. They have been reduced to the same basis by expressing the frequencies in terms of percents. Each of these curves closely approximates the normal distribution. Thus, in the case of the Illinois General Intelligence Scale, the shape of the distributions furnishes no evidence that there is a lack of discrimination.

In Figures 13, 14, and 15, we give the distributions of the achievement ages for grades III, V, VI, and VIII, for the pupils of one city. Because the original bases for translating point scores into achievement ages were found to be incorrect it is not possible to use the scores of the pupils on which the distributions for mental age given in Figure 12 are based. The distributions of achievement ages in arithmetic (Figure 13) exhibit striking departures from the normal shape in grades III and V. In grades VI and

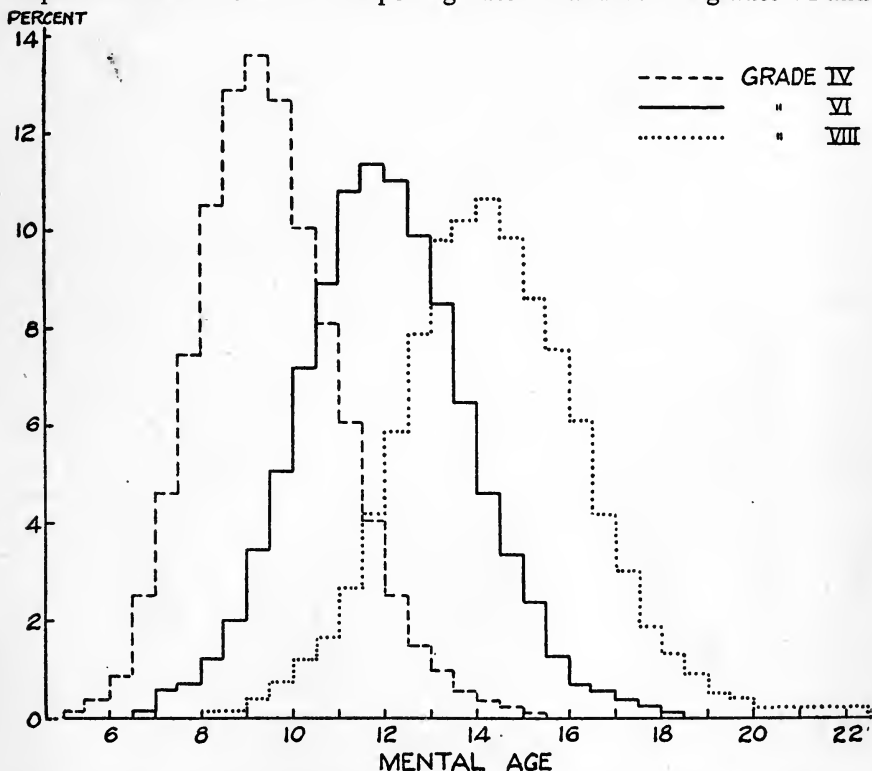


FIGURE 12. DISTRIBUTION OF PUPILS ACCORDING TO MENTAL AGE.

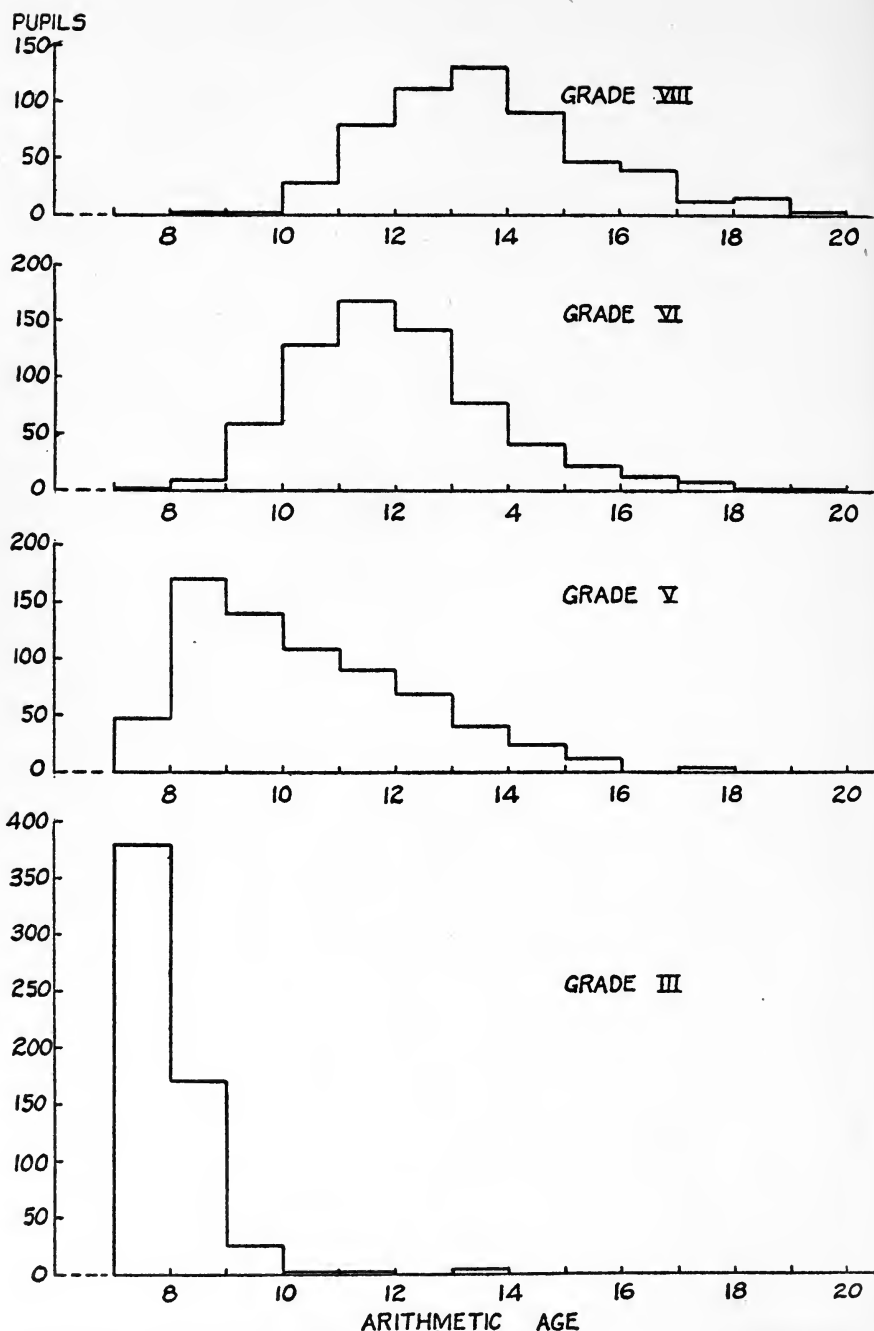


FIGURE 13. DISTRIBUTION OF PUPILS ACCORDING TO ACHIEVEMENT AGE IN ARITHMETIC.

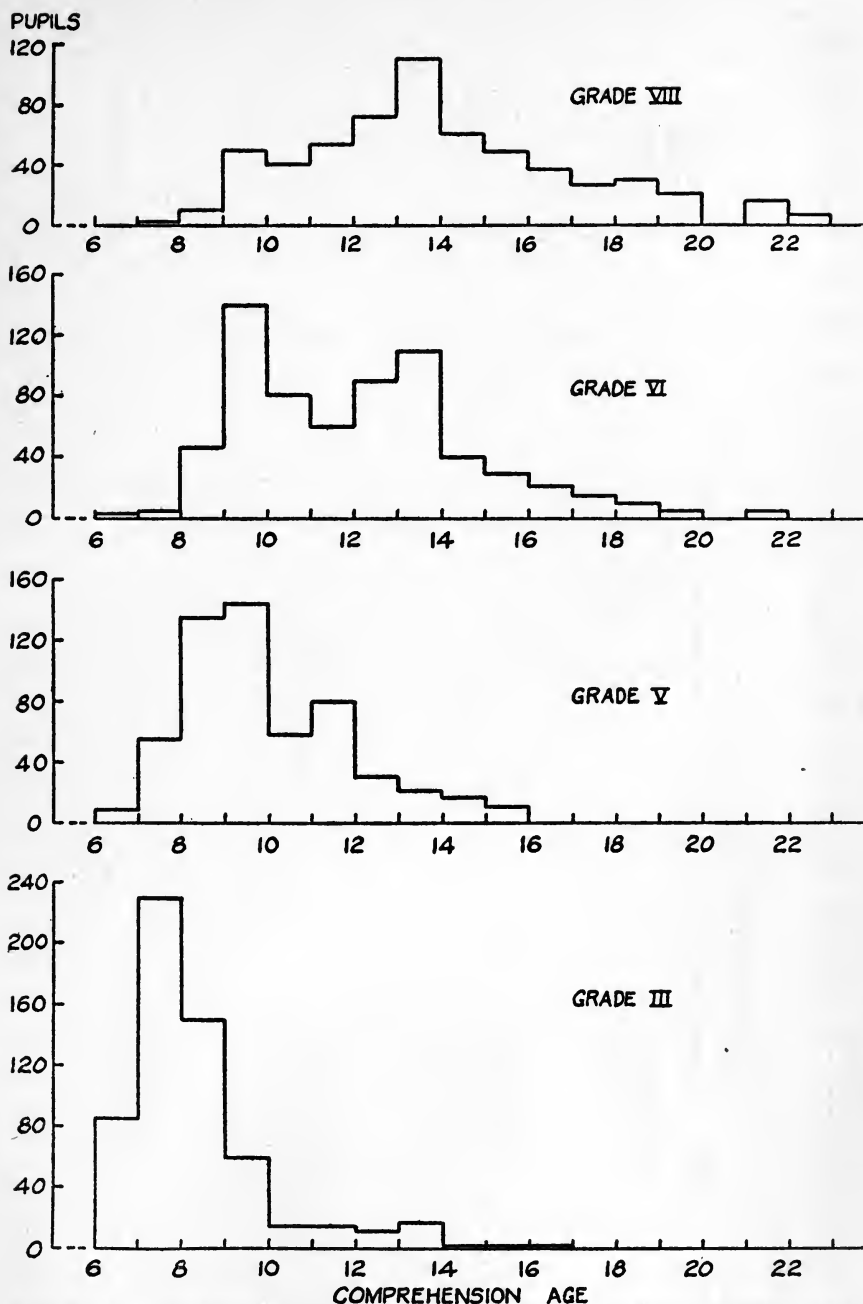


FIGURE 14. DISTRIBUTION OF PUPILS ACCORDING TO ACHIEVEMENT AGE IN COMPREHENSION OF SILENT READING.

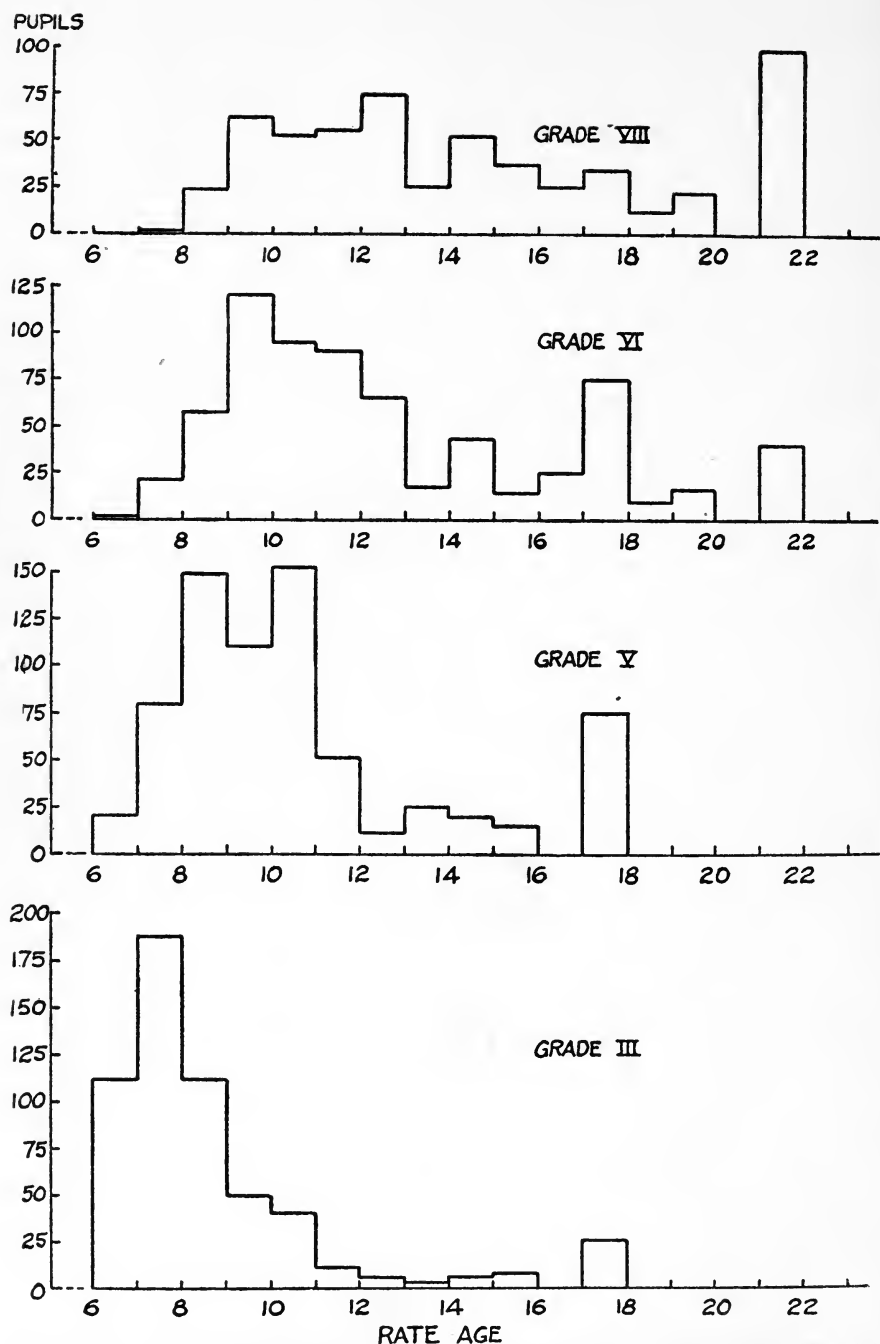


FIGURE 15. DISTRIBUTION OF PUPILS ACCORDING TO ACHIEVEMENT AGE IN RATE OF SILENT READING.

VIII the resemblance to the normal curve is very close. It will be remembered that in measuring achievement in the field of arithmetic one scale is used for grades III, IV, and V, and another for grades VI, VII, and VIII. It appears, therefore, that the scale designed for use in the lower grades does not measure accurately the arithmetical abilities of some pupils in the lower grades. The fact that such a large percent of pupils in these grades are grouped together shows that this scale fails to discriminate between some pupils having different degrees of ability.

Figure 14 gives the corresponding distributions for comprehension of silent reading. These distributions more nearly approach the normal shape. The one for the sixth grade exhibits two modes. In each grade there are a few very high scores which destroy the symmetry of the figures.

In Figure 15, we represent graphically the corresponding distributions for rate of silent reading. These graphs exhibit striking departures from the normal shape, particularly in the upper grades. It is clear, therefore, that in the case of rate of silent reading the measures must be considered lacking in accuracy in a number of cases. In the eighth grade, particularly, there is a failure to discriminate with respect to the rate of reading in the cases of many pupils who read rapidly. This tendency is also seen in the other grades. The irregularities are probably due in part to the fact that rate of reading the unconnected exercises varies from exercise to exercise. Some may be read rapidly, others very slowly.

In Figure 4 we gave the total distribution of the intelligence quotients for the pupils in grades IIIB to VIIIA, inclusive, in a large city school system. (See page 29). This distribution approximates the normal distribution very closely except for the interval from 90 to 100. The frequency in this interval is less than that for the intervals on either side. This, however, is probably due to the use of our table to calculate the I. Q.'s.

For the achievement quotients we give only the distributions for all grades combined. The distributions for the separate grade groups do not show any marked irregularities. In Figure 16, the distribution of the achievement quotients for arithmetic is given. This closely approximates the normal shape. This close approximation will also be observed in Figure 17 in the case of comprehension of silent reading. For the rate of silent reading (See Figure 18) the distribution is less symmetrical, as we might expect after noting the irregularities in Figure 15.

An examination of the distributions of the intelligence quotients and of the achievement quotients does not reveal any indication of a marked lack of discrimination except possibly in the case of the rate of silent reading. One might be inclined to conclude that there was evidence of a lack of discrimination in the case of the intelligence quotient if the distributions for mental age did not approach the normal curve as closely as they do.

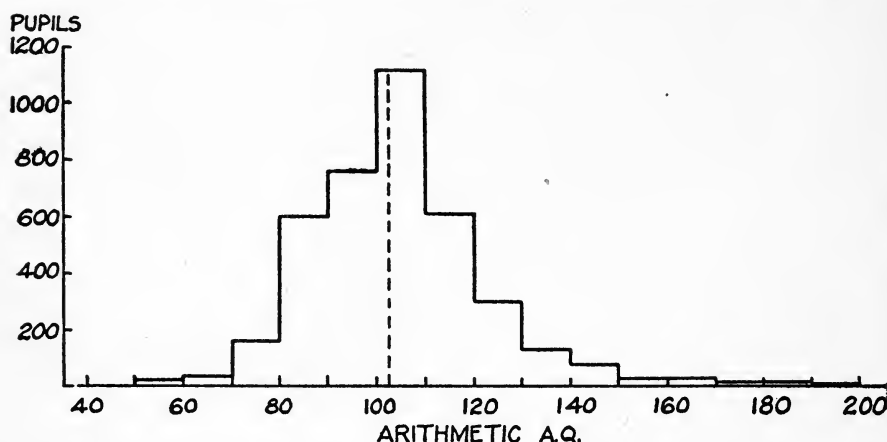


FIGURE 16. DISTRIBUTION OF PUPILS ACCORDING TO ACHIEVEMENT QUOTIENTS IN ARITHMETIC.

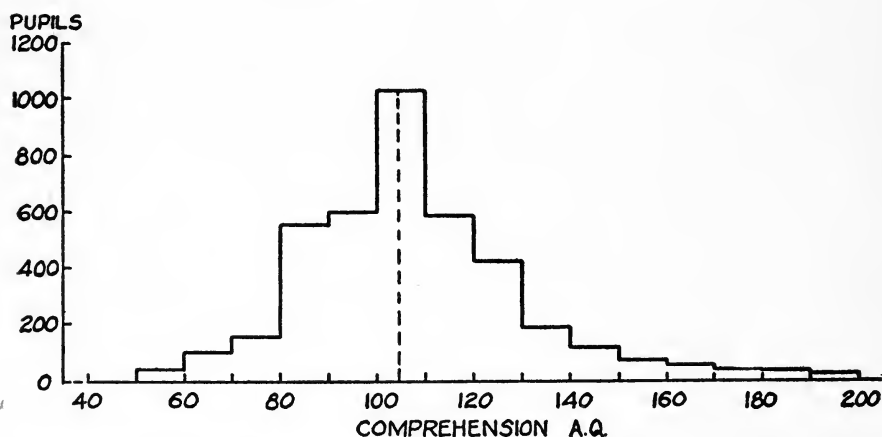


FIGURE 17. DISTRIBUTION OF PUPILS ACCORDING TO ACHIEVEMENT QUOTIENTS IN COMPREHENSION OF SILENT READING.

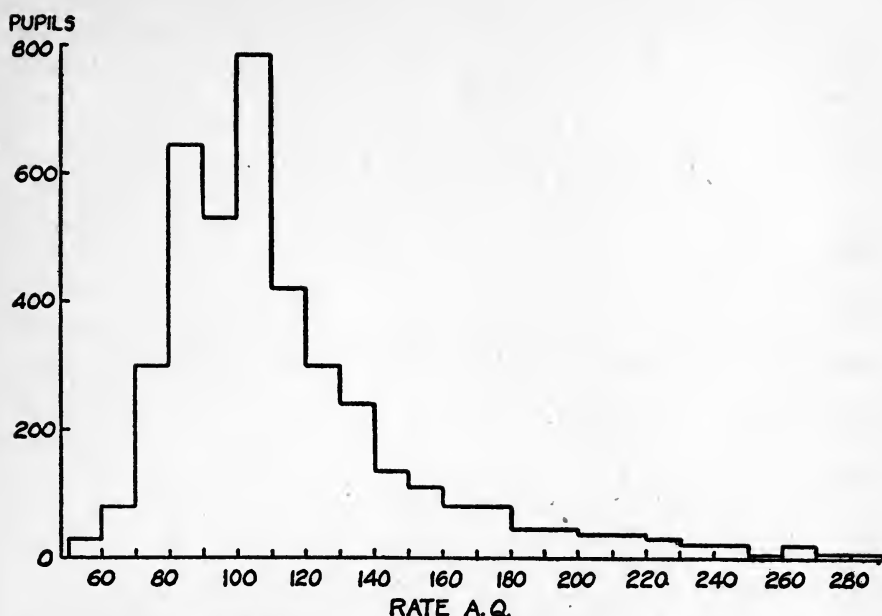


FIGURE 18. DISTRIBUTION OF PUPILS ACCORDING TO ACHIEVEMENT QUOTIENTS IN RATE OF SILENT READING.

4. **Comparison with criterion measures.** The Illinois General Intelligence Scale was given to 203 pupils whose mental ages had also been determined by the Stanford Revision of the Binet Scale for Measuring Intelligence. The correlation between the mental ages determined by these two scales is $.74 \pm .02$. The probable error of estimate is 1.2 years. This means that in 50 percent of the cases, the mental age, as determined by the Illinois General Intelligence Scale, differed from that as determined by the Binet Scale by 1.2 years or less. This lack of agreement between the measures secured by these two scales is not due solely to errors in the measures yielded by the Illinois General Intelligence Scale. The Stanford Revision of the Binet Scale for Measuring Intelligence also yields measures which involve errors of approximately the same magnitude as those of the Illinois General Intelligence Scale. (See pages 24-27).

In November, 1920, both the Illinois General Intelligence Scale, Form 1, and the National Intelligence Scale, Form 1, were given to 3615 pupils in eight elementary schools in Chicago. The correlation between the scores obtained from these two tests is indicated in Table XVI. The probable errors of estimate indicate that the agreement is not close. The probable error of estimate, when all grades are taken together, is 11.5.

This means that the departure from perfect correlation with the scores yielded by the National Intelligence Scale is greater than 11.5 points in 50 percent of the cases, and less than 11.5 in 50 percent of the cases. Since ten points are equivalent to one year of mental age, the relationship between the scores yielded by the Illinois General Intelligence Scale and by the National Intelligence Scale is approximately the same as the relationship shown to exist between the scores yielded by the Illinois General Intelligence Scale and the Stanford Revision of the Binet Scale for Measuring Intelligence.

TABLE XVI. CORRELATION BETWEEN SCORES YIELDED BY ILLINOIS GENERAL INTELLIGENCE SCALE AND BY NATIONAL INTELLIGENCE SCALE

Grade	Number of Cases	r	P.E. _{est}	P.E. _{est}
				Aver.
III A	357	0.53	9.1	0.22
IV B	416	0.70	9.6	0.18
IV A	335	0.74	8.0	0.14
V B	460	0.55	8.7	0.14
V A	285	0.47	12.0	0.19
VI B	383	0.44	12.6	0.17
VI A	259	0.67	10.8	0.13
VII B	350	0.70	11.0	0.12
VII A	210	0.68	10.3	0.11
VIII B	271	0.72	10.2	0.10
VIII A	289	0.69	10.9	0.10
All Grades	3615	0.81	11.5	0.16

The Illinois General Intelligence Scale, Form 1, was given to a number of sixth-grade pupils whose I. Q.'s, as determined for the Otis Group Intelligence Tests, were available.* The coefficient of correlation for 83 VIA pupils was $.82 \pm .02$. For 124 VIB pupils the value of r was $.83 \pm .02$. The probable error of estimate was 6.4 in the first case and 5.9 in the latter.

The Pintner Non-Language Group Intelligence Tests are represented to have a reliability coefficient of .72. This is by mental indices and not by point scores. The two sets of measures were obtained by use of the same test after an interval of two years. The number of children tested was 46. These group intelligence tests were also given during the same semester to 300 children whose mental ages had been determined by the Stanford Revision of the Binet Scale for Measuring Intelligence. The

*The writer is indebted to Superintendent L. W. Keeler, Michigan City, Indiana for these data.

coefficient of correlation between the point scores yielded by the Pintner Group Test and the mental ages determined by the Binet Test was .80.*

No data are available at this time for making comparison between the measures of achievement yielded by the achievement scales included in the Illinois Examination and by other similar scales. Neither are data available for comparison of measures of achievement with teachers' estimates.

5. Inferences concerning validity based upon the structure of the test and its administration. In the case of the Illinois General Intelligence Scale the sub-tests have frequently been used by other makers of instruments for measuring general intelligence. At the time the Illinois General Intelligence Scale was constructed a number of other intelligence scales were analyzed with reference to sub-tests and the ones most frequently found were incorporated in this scale.

The intercorrelations between the seven sub-tests were studied by choosing at random two sets of 120 test papers each. In securing these samples ten papers were chosen from each half grade. The correlation of each sub-test with each other sub-test and with the total test score is given in Table XVII. The upper number in each case refers to the first sample and the lower to the second sample. It will be noted that the correlation of a sub-test with the total score is higher than the correlation with the other sub-tests. The correlations with the total score are high in every case. Sub-tests 2 and 3 have the highest coefficients of correlation. The intercorrelations between the sub-tests are relatively high but it must be remembered that we have used a wide range of talent extending from IIIB to VIIIA. The table, however, is evidence that the overlapping between the sub-tests is not great and is approximately uniform.

It is interesting to note the differences between the corresponding coefficients yielded by the two samples. The differences are never large and are within the range indicated by the probable errors of the coefficients of correlation due to sampling.

The Illinois General Intelligence Scale is explicitly a verbal test. Ability to read is a prerequisite. For this reason it may be urged that it does not permit non-verbal elements of intelligence to function. This objection, of course, applies to other verbal tests. No data are at hand to show the limitation which this feature places upon the scale.

The silent reading test included is a revision of Monroe's Standardized Silent Reading Test. In this revised form certain features of the original test which were found unsatisfactory have been eliminated. The scoring has been made objective. The exercises are more uniform and a more

*Pintner, Rudolph and Marshall, Helen. "Combined mental-educational survey." *Journal of Educational Psychology*. 12:32-43, January, 1921.

TABLE XVII. INTERCORRELATIONS BETWEEN THE SUB-TESTS OF THE ILLINOIS GENERAL INTELLIGENCE SCALE, FORM I*

Sub-Tests	Sub-Tests							
	1	2	3	4	5	6	7	Total
1		.58	.63	.51	.51	.57	.56	.77
		.68	.73	.48	.44	.60	.62	.84
2	.58		.74	.65	.49	.67	.64	.83
	.68		.75	.60	.61	.73	.70	.89
3	.63	.74		.63	.57	.64	.68	.88
	.73	.75		.61	.70	.72	.76	.93
4	.51	.65	.63		.41	.58	.52	.77
	.48	.60	.61		.50	.55	.54	.78
5	.51	.49	.57	.41		.51	.47	.70
	.44	.61	.70	.50		.51	.57	.77
6	.57	.67	.64	.58	.51		.55	.79
	.60	.73	.72	.55	.51		.61	.80
7	.56	.64	.68	.52	.47	.55		.79
	.62	.70	.76	.54	.57	.61		.86
Total Score	.77	.83	.88	.77	.70	.79	.79	
	.84	.89	.93	.78	.77	.80	.86	

*Two coefficients are given for each pair. Each was calculated from a random selection of ten test papers from each half grade group III B to VIII A, inclusive.

precise measurement of the rate of silent reading is secured. Experience has shown that the test is too short for the time limit allowed. This, in the case of the most fluent readers, prevents one from securing valid measures of reading ability. To one acquainted with the nature of reading ability and its measurement, this one test obviously measures in a general way, only one type of silent reading ability. To measure completely all phases of silent reading ability would require a battery of tests.

In the case of Monroe's General Survey Scale in Arithmetic the sub-tests are judged to represent the most important types of examples learned by pupils in the sequence of grades for which they are intended. The choice represents the judgment of the author but is not inconsistent with other groups of tests used to measure ability of pupils in the operations of arithmetic. A single general score has been used to describe the pupil's

ability. This is, obviously, a composite not only of the scores of the different sub-tests but also of the dimensions of rate and accuracy. However, since a single general measure is desired this does not constitute a serious criticism of the scale.

Summary for validity. By way of summary we may say that the scales which make up the Illinois Examination compare favorably in respect to validity with our best tests. It is, however, clear that the scales possess certain limitations which should be kept in mind when the scores are interpreted.

VI. VALIDITY OF SIGNIFICANCE

Probably all readers will admit that the abilities measured by this battery of tests are important in the education of children. They are probably the abilities most often measured. The Illinois Examination includes a plan, however, for combining measures of intelligence and measures of achievement. It is, therefore, appropriate that we point out the significance of the comparison proposed.

Significance of achievement quotients. In Figure 19, the achievement ages for both arithmetic and reading average are shown for City A and City B. It is clearly evident, from this figure, that when considered

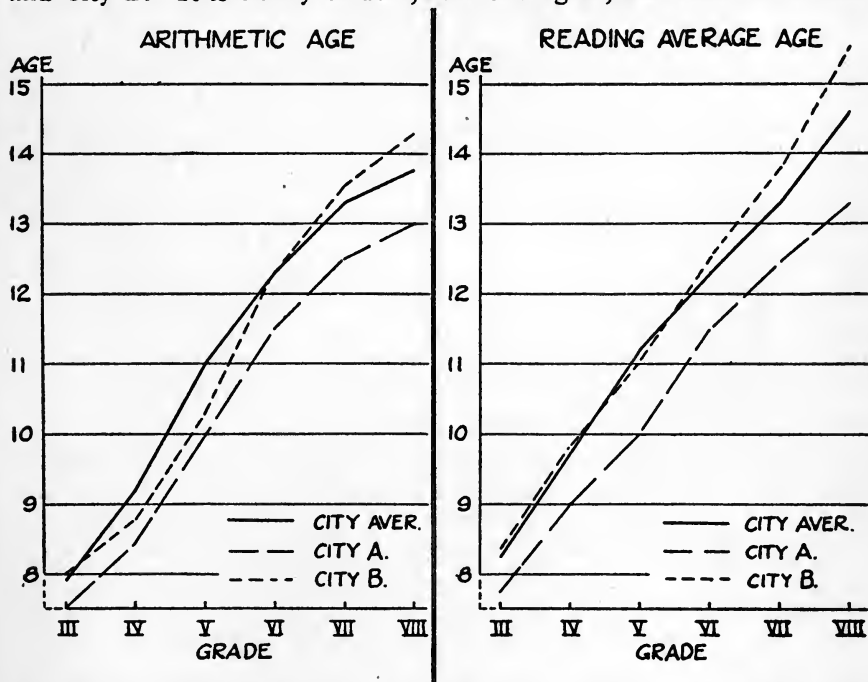


FIGURE 19. ACHIEVEMENT AGES FOR ARITHMETIC AND FOR READING (AVERAGE) FOR CITY A AND CITY B.

grade for grade, City B is distinctly superior to City A in both arithmetic and reading. The superiority is more pronounced in the case of reading. If this were the only information at hand for these two cities the conclusion would be that the schools of City B were superior to those of City A, and that City A was below the city average.

In Figure 20, the median mental age and the median I. Q. for the several grades of these two cities are represented. The school system in City B is shown to have been so organized that in each grade the median mental age of the pupils is about one year greater than in City A. The median intelligence quotients are also distinctly higher. In fact, City B, except in the third grade, is distinctly above the city average. These facts mean that, grade for grade, City B has superior pupil material. This is probably due to a difference in the organization of the school systems of the two cities. In City B, the school system is so organized that in the upper grades the pupils are highly selected. In City A, the opposite condition prevails. Therefore, we should expect City B to secure distinctly superior achievements, grade for grade.

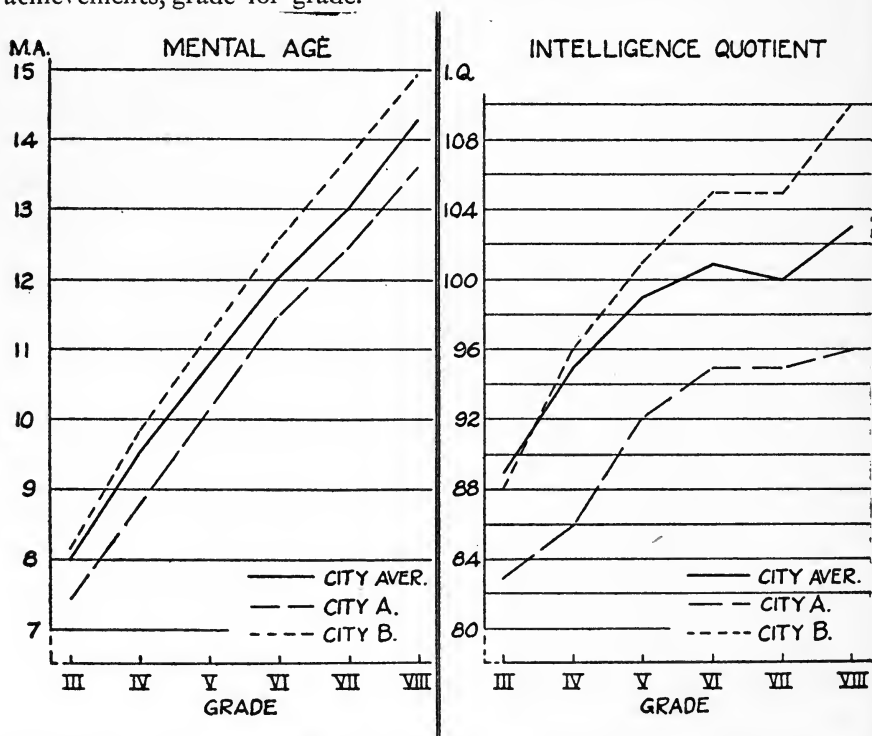


FIGURE 20. MEDIAN MENTAL AGE AND MEDIAN I. Q. FOR EACH GRADE FOR CITY A AND CITY B.

We now have to answer the question, "Is the school system of City B as effective as that of City A when the character of the pupil material in the two cities is taken into account." Figure 21 answers this question. It shows the achievement quotients for these two cities. In the case of arithmetic the quotients for City B are, except in the eighth grade, below those for City A. They are also below the city average. In the case of reading City B surpasses City A only in the seventh and eighth grades. Therefore, we must conclude that, with the exception of the eighth grade and in part the seventh, the school system of City B is less effective in the teaching of the operations of arithmetic and silent reading than the school system of City A. By using the mental ages of the pupils we have thus been able to avoid an erroneous interpretation of the measures of achievement for these two cities.

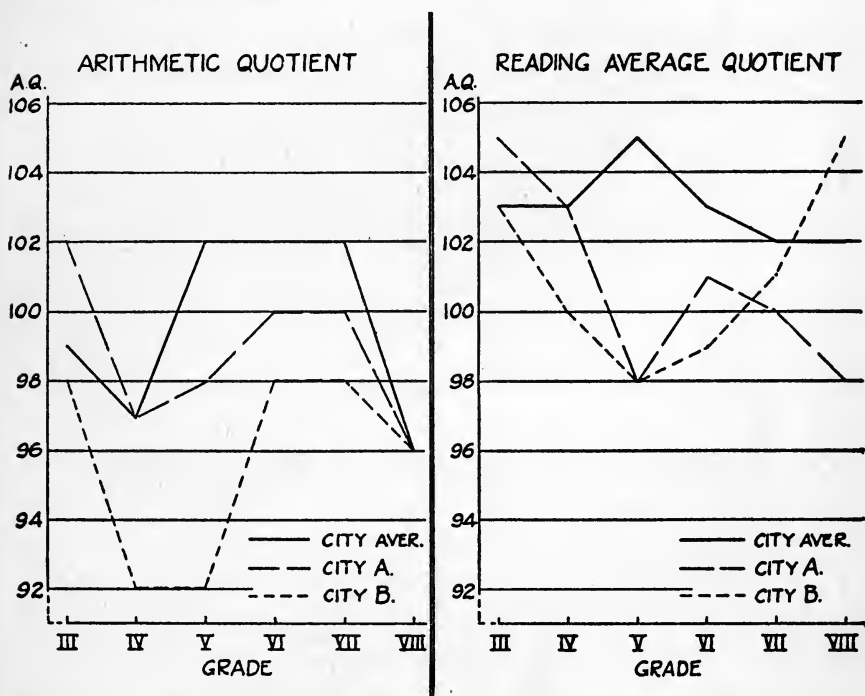


FIGURE 21. ACHIEVEMENT QUOTIENTS FOR EACH GRADE FOR CITY A AND CITY B.

A pupil's mental age constitutes an individual norm with which his achievement ages may be compared. The results of this comparison are expressed by the achievement quotients. The value of individual norms and the achievement quotient are illustrated by Figure 22. This figure

A scatter plot showing the relationship between 'Standard' (x-axis) and 'A.Q.' (y-axis). The x-axis ranges from 4 to 16, and the y-axis ranges from 50 to 200. The data points are as follows:

Standard	A.Q.
4	108
5	70
6	88
7	83
8	88
8	145
8	178
9	68
9	83
9	119
9	140
10	82
10	103
10	104
10	105
10	106
10	107
10	108
10	109
10	110
10	111
10	112
10	113
10	114
10	115
10	116
10	117
10	118
10	119
10	120
10	121
10	122
10	123
10	124
10	125
10	126
10	127
10	128
10	129
10	130
10	131
10	132
10	133
10	134
10	135
10	136
10	137
10	138
10	139
10	140
10	141
10	142
10	143
10	144
10	145
10	146
10	147
10	148
10	149
10	150
10	151
10	152
10	153
10	154
10	155
10	156
10	157
10	158
10	159
10	160
10	161
10	162
10	163
10	164
10	165
10	166
10	167
10	168
10	169
10	170
10	171
10	172
10	173
10	174
10	175
10	176
10	177
10	178
10	179
10	180
10	181
10	182
10	183
10	184
10	185
10	186
10	187
10	188
10	189
10	190
10	191
10	192
10	193
10	194
10	195
10	196
10	197
10	198
10	199
10	200
11	90
11	152
12	120
13	118
13	167
14	106
15	123
16	133

The lowest score in the class is four. If we had only the grade norm, all we could say about this pupil would be that he is conspicuously below standard and at the foot of his class. His A. Q. shows that in comparison with his own norm he has achieved more than is usually achieved by a pupil of his mentality. In fact when his mental age is considered, he is one of the "good" pupils in his class.

64

lengths of time. Since the achievement ages are expressed in the same units, they may be combined as point scores cannot be combined. For example, we have no obvious way of expressing the total achievement of a child who scores 155 in rate, 11 in comprehension, and 49 in arithmetic. By reference to Table XI, however, we observe that these point scores indicate, respectively, achievement ages of 12-6, 11-6, and 10-6 years. On the assumption that the abilities to which the point scores refer are all equally important, we may obtain the simple average of these ages, and we may thus express the composite achievement age as 11-6 years.

The translation of the point scores into age scores makes the measures more easily interpreted. Even without access to grade norms one is able to interpret partially a mental age of 13 or an achievement age of 9 years, 6 months. The quotients are also easily interpreted because in every case the norm is approximately 100.

The grade norms in the case of the achievement test do furnish appropriate educational objectives. The norms for the intelligence scale cannot, of course, be considered objectives because of the assumption that traits measured by this scale are not covered by school instruction.

VII. NORMS

For the Illinois Examination two types of norms are available: age norms and grade norms. The Illinois General Intelligence Scale has been standardized with respect to chronological age and the achievement tests with respect to mental age. Both the age scores and quotients have been standardized with respect to school grade. The grade norms are given in Table XVIII. The age norms are incorporated in the rules for translating the point scores into age scores. In both cases the norms are for *first trial* scores only. When the tests are repeated these norms should not be used. (See page 68). The grade norms for both ages and quotients are given in Table XVIII for three population groups; rural schools, city schools, and a general group, including city, town, and rural schools. The median scores for any grade or any city may be compared with these grade norms. It is helpful to make this comparison by expressing the median scores as deviations from these norms. When the median score is larger than the corresponding grade norm the deviation is positive, when less, negative. In Table XIX the median scores for a city are expressed as deviations from the grade norms. In the case of ages, the deviations are in terms of months: -6 indicates that the median age for this city is a half year below the grade norm, +8 that it is eight months above the grade norm.

TABLE XVIII. GRADE NORMS BASED UPON 55,000 OCTOBER SCORES*

Grade		Ages					Quotients				
		Mental	Arithmetic	Average	Silent Reading		Intelligence	Arithmetic	Average	Silent Reading	
					Comprehension	Rate				Comprehension	Rate
III	Rural Cities General	7-10	8-0	7-11	7-11	7-11	90	102	101	101	101
		7-11	7-10	8-2	8-2	8-2	89	99	103	103	103
		7-10	8-0	7-11	7-11	7-11	90	102	101	101	101
IV	Rural Cities General	9-4	9-3	9-1	9-0	9-1	96	99	97	96	97
		9-5	9-2	9-8	9-9	9-7	95	97	103	103	102
		9-4	9-2	9-5	9-4	9-5	95	98	101	100	101
V	Rural Cities General	10-6	10-11	10-1	10-0	10-2	98	104	96	95	97
		10-8	10-11	11-2	11-6	10-10	99	102	105	108	102
		10-8	10-11	10-8	10-9	10-7	98	102	100	101	99
VI	Rural Cities General	11-9	12-0	11-4	11-1	11-7	100	102	96	94	99
		12-0	12-3	12-4	12-6	12-2	101	102	103	104	101
		11-11	12-0	12-0	12-0	11-11	101	101	101	101	100
VII	Rural Cities General	13-1	13-2	13-1	13-1	13-0	101	100	100	100	100
		13-1	13-6	13-5	13-5	13-4	101	102	101	101	101
		13-1	13-4	13-4	13-3	13-2	101	101	101	100	101
VIII	Rural Cities General	14-3	14-3	14-2	13-10	14-6	106	103	100	97	102
		14-4	14-4	14-11	14-6	15-10	104	100	102	100	106
		14-3	14-2	14-7	14-2	14-10	104	101	101	98	104

*These are first trial scores.

TABLE XIX. MEDIAN SCORES FOR A CITY EXPRESSED AS DEVIATIONS FROM THE GRADE NORMS

Grade	Ages					Quotients				
	M.A.	Arith.	Aver.	Reading		I.Q.	Arith.	Aver.	Reading	
				Comp	Rate				Comp	Rate
III	-1	-2	-1	+4	-6	+2	-6	-1	+2	+0
IV	0	-3	0	-1	0	+2	-1	-2	+2	-4
V	-2	-6	-2	-3	0	+2	-5	+4	+4	+5
VI	+3	-9	-1	+8	-7	+6	-7	-1	-1	-2
VII	-2	-1	+4	+10	-1	0	+1	-2	0	-3
VIII	+4	+2	+1	+8	-7	0	+5	-1	+1	-3

The intelligence quotients derived from the Illinois General Intelligence Scale exhibit a greater degree of variability than those derived from the Binet Scale for Measuring Intelligence. For this reason one must use a different basis for interpreting the I. Q. 's in terms of degrees of brightness. Intelligence quotients derived from the Illinois General Intelligence Scale are estimated to be distributed as shown below* The appropriate interpretation is indicated in the left-hand column.

Degree of brightness	I.Q.	Percent of all children included
"Near" genius or genius.....	140 and above	1
Very superior.....	125-139	6
Superior.....	115-124	13
Normal or average.....	85-114	60
Dull.....	75-84	13
Border-line.....	60-74	6
Feeble-minded.....	Below 60	1

Achievement quotients exhibit a somewhat greater degree of variability than intelligence quotients. Their distribution differs from the normal by showing a greater degree of variability above the median (approximately 100) than below the median. It is, however, possible to divide the distribution so that the percent of pupils included in each division corresponds to that used for the interpretation of intelligence quotients. The following scheme for interpretation is suggested:

*It is necessary to estimate this distribution because the original rule for translating point scores into mental ages was found to be unsatisfactory.

Quality of pupils' achievement	Achievement Quotient	Percent of pupils included
Very superior.....	165 and above	1
	135-164	6
Superior.....	117-134	13
Average.....	83-116	60
Poor.....	71-82	13
	55-70	6
Failure.....	Below 55	1

Equivalence of Form 1 and Form 2. In considering norms for the two forms of the Illinois Examination it is necessary to inquire concerning the equivalence of the two forms. In case one form is easier than the other the same norms cannot be used. Data with reference to the equivalence of the scales which make up the Illinois Examination have already been given (See pages 9-18). These facts may be summarized and expressed in the following form which is more convenient to use. In order to reduce Form 2 age scores to the equivalence of Form 1 scores multiply by the following correction numbers:

	Correction Number
Illinois General Intelligence Scale.....	.98
Monroe's General Survey Scale in Arithmetic, Scale I.....	.96
Monroe's General Survey Scale in Arithmetic, Scale II.....	1.00
Monroe's Standardized Silent Reading Test I, Revised, Comprehension.....	1.00
Rate.....	1.05
Monroe's Standardized Silent Reading Test II, Revised, Comprehension.....	.99
Rate.....	.97

Practise effect when a test is repeated. The grade norms given in Table XVIII are for the first application of the Illinois Examination. When it is given a second time pupils will tend to make higher scores because of their acquaintance with the nature of the tests. The amount of increase varies. If pupils are "coached" upon the tests a large increase is to be expected. When a period of several months intervenes between the first and second trials and the pupils have received no training upon the exercises of the tests, the increase appears to be small and in some cases can be neglected without serious error. No evidence was obtained relative to the effect of using the same form instead of a different form.

In order to ascertain the effect of practise when the second application immediately follows the first, both Form 1 and Form 2 were given to a number of pupils. Due to a miscarriage of plans the practise effect for the silent reading tests was not determined. For the Illinois General Intelligence Scale the average practise effect is approximately 5.0 points, or six months of mental age, if the returns from the eighth grade pupils are not used. In this grade unusual conditions appear to have prevailed and when the scores from it are included the practice effect is approximately 7.0 points. For Monroe's General Survey Scale in Arithmetic the average practise effect is approximately 3.2 points in grades III to V and 4.5 points in grades VI to VIII. It is, therefore, obvious that when the Illinois Examination is repeated after only a short interval the second trial scores must be corrected before they can be compared with those obtained from the first trial.

In one school the teachers of 134 pupils gave special drill and instruction to their pupils after Form 1 of the Illinois Examination had been given in November. The teachers did not know that Form 2 was to be given later, and did not have in mind, therefore, preparing the pupils for it. Their instruction was, to a slight extent, based upon Form 1 test papers. Believing that their pupils were rather weak in knowledge of vocabulary and in synonym-antonym, special drill was given along these lines in language work. In arithmetic practise was given upon those combinations where the pupils seemed weak. In reading there was some special drill for increasing the rate of silent reading. The gains made by the pupils under these teachers are given in Table XX.

It is commonly assumed that general intelligence is unaffected by school instruction. A median gain of 43.2 points or slightly more than four years in mental age within a period of six months indicates that mental age, as measured by such an instrument as the Illinois Intelligence Scale, is affected by class room instruction. It is, therefore, necessary to exercise a great deal of caution in interpreting changes in the intelligence scores derived from two successive testings separated by a considerable time interval. Even in the case of a first trial the scores obtained will be misleading if the pupils have received any special preparation for the test.

One would naturally expect that the achievement scores in arithmetic and silent reading would be materially affected by instruction. Table XX shows very large gains in achievement. The increases in achievement quotients show that except for the comprehension of silent reading the gains are relatively larger in achievement than in mental age. It is not unlikely that some of the increases in the May scores over the November scores are due to the pupils being more familiar with the testing procedure. The effect of familiarity was not investigated but due caution

TABLE XX. GAINS DUE TO SPECIAL INSTRUCTION UPON THE ILLINOIS EXAMINATION

	Median Point Scores			Median Quotients		
	Nov. 1920	May 1921	Gain	Nov. 1920	May 1921	Gain
Intelligence.....	57.6	100.8	43.2	99.6	128.0	28.4
Arithmetic.....	48.2	120.0	71.8	114.0	137.4	23.4
Comprehension..	10.9	15.5	4.6	105.1	102.0	-3.1
Rate.....	171.4	237.9	66.5	122.0	139.4	17.4

should be exercised in interpreting the gains in achievement as being the result of instruction directed to the needs of the pupils.

